Techniques for Cosmological Analysis of Next Generation Low to Mid-Frequency Radio Data

UNIVERSITY of PORTSMOUTH

Michael Tarr

Department of Technology

THE THESIS IS SUBMITED IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE AWARD OF THE DEGREE OF DOCTOR OF PHILOSOPHY OF THE UNIVERSITY OF PORTSMOUTH

May 2018

Declaration

Whilst registered as a candidate for the above degree, I have not been registered for any other research award. The results and conclusions embodied in this thesis are the work of the named candidate and have not been submitted for any other academic award.

This dissertation contains 51,523 words not including appendices, bibliography, footnotes, tables and equations, and has 68 figures.

Michael Tarr
May 2018

Acknowledgements

First, I would like to give my most sincere thanks to David. It goes without saying that I can not imagine having completed this work without you. Your enthusiasm and belief in me has been a constant source of motivation from day one to the eleventh hour. You are an inspiration and model of a perfect superior. I cannot overstate how much I appchiate your effort, and I am eternally grateful.

My love and thanks also to my parents, who have provided nothing but loving support, despite all my whims and phases. I can only hope it was all worth it to find something I have finally stuck with.

Thanks should also go to Xan and Matthew, for taking precious time away from their own research to indulge my pet machine learning project. I owe you both and hope to always be friends.

Rebecca[1], save the final thanks for you. Your support and patience has been wonderful, I will never forget it. I promise to save you from the washing, and make up for all the lost holiday! All my love, "*Always*".

[1] *Hello to Jason Isaacs.*

Dissemination

Publications

Tarr, M. and Bacon, D. "Measuring gravitational weak lensing from radio visibilities" (in prep).

Tarr, M., Sabater, J. et al., "Spectra of faint radio sources in the ELAIS-N1 field" (in prep).

Smith, D. J. B. et al., including **Tarr, M.** "The WEAVE-LOFAR Survey, SF2A-2016: Proc. of FSAA, arXiv:1611.02706 [astro-ph.CO].

Recent Presentations

Nov. 2016, Science for the SKA Generation: *"Radio Weak Lensing: Potential for the SKA"*

Jun 2016, UCL seminar: *"Radio Weak Lensing: Exploring methods for shear estimation"*

Apr. 2016, SKA – Delivering the Science: *"Techniques and tools for weak lensing with the SKA"*

Mar. 2016, WEAVE-LOFAR team meeting: *"LOFAR-BOSS ELIAS-N1: SDSS-III observations of radio galaxies"*

Sept. 2015, ICG-KASI meeting, Portsmouth: *"Radio Weak Lensing: and direct shear mapping using F.I.L.M"*

Jul. 2015, National Astronomy Meeting, Llandudno: *"Fourier Inspection of Lensing Modes (F.I.L.M): A new weak lensing estimator for next generation of radio telescopes"*

Abstract

For centuries, astronomical instrumentation has continued to develop and improve. This has led to the incredible precision seen in recent observations, with ever greater accuracy to come. From their measurements, cosmologists have produced a complex model of the Universe. This indicates that most of the energy in the Universe is in some "dark" form that does not interact directly with the electromagnetic spectrum. One of the greatest challenges of modern cosmology is the understanding of these dark components. In particular weak gravitational lensing has emerged as powerful tool for probing these parts of the cosmological model.

An area of observation which has undergone an accelerated improvement in recent years, is radio astronomy. The radio regime therefore, is an exciting source of new discoveries and further cosmological constraints. Several studies have shown that the lensing signal can be detected at radio wavelengths[2], while others have demonstrated the significant impact of combining the results from next-generation radio and optical telescopes[3]. In order to fully realise the potential of lensing and other measurements for cosmology, a comprehensive understanding of the radio galaxy population is also required.

In this thesis, I consider the future of radio observations for precision cosmology and the required developments in analysis techniques. I aim to contribute to this progress by focusing specifically on: (i) Providing novel estimators for radio weak lensing measurements and (ii) Studying the redshifts and statistics of galaxy populations in the radio, using optical spectroscopy. Both of these areas will soon be the subject of large next-generation observations, in the form of the Square Kilometre Array (SKA[4]) continuum surveys; and a joint experiment between a next-generation spectroscopy facility (WEAVE[5]) for the William Herschel Telescope (WHT[6]) and the LOw Frequency ARray (LOFAR[7]).

[2]Chang et al. (2004); Patel et al. (2010)
[3]Bonaldi et al. (2016); Camera et al. (2017); Harrison et al. (2016)
[4]https://skatelescope.org/
[5]http://www.ing.iac.es/weave/about.html
[6]http://www.ing.iac.es/Astronomy/telescopes/wht/
[7]http://www.lofar.org/

Table of contents

List of figures

List of tables

Chapter 1

Cosmology

For centuries the big cosmological questions of the origin, form and fate of the Universe have inspired people to push the limits of technology and accepted knowledge. The desire to accurately measure the transit of Venus in 1761, for example, led to an astounding international collaboration of scientists at a time of intense global conflict. Modern cosmology is concerned with problems on a far larger scale than the size of the Solar System, and the practice of international collaboration is just as important today as it was in the 18th Century.

With almost instantaneous communication around the world, a slew of ever more powerful instruments, and gravitational waves opening completely new window on the Universe, the next decade looks to be the golden age of cosmology. These developments can only reach their full potential if the analytic tools we use keep pace with the data. As such, this thesis considers the future of radio observations, an area in which data quality has exploded in recent years, for the future of cosmology. In the next six chapters I will examine some of the most promising next generation science cases and attempt to contribute to their success in the spirit of the transit chasers from over a quarter of a millennium ago. I begin by surveying the current state of cosmology.

The modern consensus in cosmology has developed from the basis of the cosmological principle. This removes the observer from any privileged position in the Universe, allowing theoretical models to then be tested against data for inconsistencies with this ideal. Currently, the most widely accepted model for the Universe is that of a flat (but perturbed) 4-dimensional space-time where structure has formed from early over-densities, and the collapse of cold dark matter (CDM). This, together with an accelerating expansion due to a cosmological constant (Λ), can reproduce most features seen in recent astronomical observations. Thus, despite the inclusion of two unidentified and unobserved components, and no explanation for

the absence of a large "vacuum energy" expected from simple particle physics calculations, ΛCDM remains the best theory of the Universe so far.

In this Chapter, I will discuss the theoretical and observational evidence for the ΛCDM model, and how a universe of this kind can be measured and constrained. I will also consider what implications this has for our Universe, and the various theories for the nature of dark matter and dark energy. Finally, I include a description of the current probes that are being used to constrain cosmological parameters.

1.1 Basis of Theoretical Cosmology

This section explores the theoretical underpinning of the current cosmological theory. I investigate the fundamental principles that have led ΛCDM to become the concordance model of cosmology, and the predictions and implications this has for our Universe.

From the very beginning we will assume the cosmological principle, along with General Relativity, as the foundation from which to build our model. This principle demands that the Universe obey two simple yet powerful properties, isotropy and homogeneity.

Isotropy: This claims that there is no preferred direction in the Universe. When averaging over large enough scales observations in any direction will be statistically concordant.

Homogeneity: This claims that there is no privileged location in the Universe. On large enough scales one region is equivalent to another, and observers at any location all experience the same global history.

These postulates have to some degree been confirmed through observations of the Cosmic Microwave Background (CMB) and galaxy surveys (see section 1.4 for more details).

1.1.1 General Relativity

After it was proposed in 1686, Isaac Newton‘s theory of gravity was used to successfully predict the movements of many celestial bodies and precisely explain the orbits of planets in the Solar System for over two hundred years. However, in 1846 a French Astronomer, Urbain Le Verrier, discovered that over the course of a century Mercury‘s orbit had drifted slightly

compared with what would have been anticipated from Newton's theory. This inconsistency remained unanswered until the early nineteenth century, when Albert Einstein developed the theory of General Relativity (GR). This marked a significant step forward in humanity's understanding of the Universe, explaining that gravity actually results from the curvature of time and space by mass rather than some mysterious force.

In the 4-dimensional manifold that GR uses to represent space-time, a metric tensor is needed to describe the distance between two points since the geometry is no longer Euclidean. Generically, an infinitesimal distance ds is described using the metric tensor $g_{\mu\nu}$, assuming the Einstein summation convention, as,

$$ds^2 = g_{\mu\nu} dx^\mu dx^\nu \tag{1.1}$$

Where μ and ν are the four components describing one time (dx^0) and three spatial ($dx^{1,2,3}$) coordinates.

The Einstein field equations illustrate and describe the interaction of matter and curvature in GR. The equations consist of the Einstein tensor ($G_{\mu\nu}$, a description of geometry), $T_{\mu\nu}$ the stress-energy tensor containing information about matter and energy, and the metric tensor in equation 1.1. They are conventionally written as (D'Inverno, 1992),

$$G_{\mu\nu} = \frac{8\pi G}{c^4} T^{\mu\nu} + \Lambda g_{\mu\nu} \tag{1.2}$$

The Λ in the second term of the field equations is called the "cosmological constant". This was introduced by Einstein to counter the contracting effect of gravity which contradicted his preference of a steady state universe. Einstein removed this term when Hubble presented evidence that the Universe is expanding (see section 1.2.1), famously referring to it as his "biggest blunder". However, Λ has been reintroduced into the concordance model as a possible solution for the accelerated expansion rate seen in the recent Universe.

1.1.2 Friedmann–Lemaître–Robertson–Walker Cosmology

By taking only the most basic requirements of the cosmological principle, those of isotropy and homogeneity, together with the Einstein Field equations, Robertson and Walker constructed a metric which can be used for testing against observational data in cosmology. In order to satisfy isotropy we simply demand spherical symmetry. Using spherical coordinates this means distance is independent of direction given by the angles θ and ϕ. Allowing the

origin of the coordinates to be arbitrary and restricting curvature to a constant value ensures homogeneity. This results in the following general metric as given by, e.g. D'Inverno (1992),

$$ds^2 = -c^2dt^2 + a^2(t)\left(\frac{1}{1-Kr^2}dr^2 + r^2(d\theta^2 + \sin^2\theta d\phi)\right) \tag{1.3}$$

K is the curvature of the Universe which recent measurements have found to be very close to zero as discussed in 1.4. We call a model which conforms to this metric a "Robertson-Walker" (RW) space-time. A universe defined by an RW metric with $K = 0$ can be described as having a flat or euclidean geometry.

By solving the Einstein equations (1.2) using the metric above, we can describe how space-time and matter interact in an isotropic and homogeneous Universe. Firstly, we make the simplifying assumption that all matter in the Universe can be described by a perfectly homogeneous fluid with spatially uniform density ($\rho(t)$) and pressure ($p(t)$). This reduces the stress-energy tensor to,

$$T_{\mu\nu} = \begin{pmatrix} \rho(t) & 0 & 0 & 0 \\ 0 & p(t) & 0 & 0 \\ 0 & 0 & p(t) & 0 \\ 0 & 0 & 0 & p(t) \end{pmatrix} \tag{1.4}$$

The field equations can now be solved and result in Friedmann's equation,

$$H^2 = \left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho + \frac{\Lambda}{3} - Kc^2a^{-2} \tag{1.5}$$

and the acceleration equation,

$$\frac{\ddot{a}}{a} = -\frac{4\pi}{3}G\left(\rho + 3pc^{-2}\right) + \frac{\Lambda}{3} \tag{1.6}$$

The fluid equation can then be obtained from the difference of the acceleration equation with the time derivative of Friedmann's equation,

$$\dot{\rho} + 3\frac{\dot{a}}{a}(p+\rho) = 0 \tag{1.7}$$

A model using a Robertson-Walker metric that also satisfies the relations derived by Freidmann is called a Friedmann–Lemaître–Robertson–Walker (FRW) metric. Specifying a relationship or an equation of state such that $w = p/\rho = constant$, we can show how the evolution of the Universe is predicted by FRW models. If the Universe is flat, i.e. $K = 0$ in equation 1.3, then density will evolve inversely with the expansion rate as,

$$\rho \propto a^{-3(1+w)} \tag{1.8}$$

1.2 Observational Cosmology

In this section I explore the way in which data has shaped our theories of the Universe and has challenged and constrained our models for cosmology. I also consider how we can describe and measure observations to further our understanding.

1.2.1 Hubble's Law

The discovery by Hubble (1929) that galaxies are receding at a rate proportional to their distance ($\mathbf{v} \propto \mathbf{r}$), and thus that the Universe is expanding was perhaps the biggest milestone in cosmology since the move from Newtonian physics to Relativity. Given that this velocity is in the direction of galaxies i.e. $\hat{\mathbf{v}} = \hat{\mathbf{r}}$ this relationship can alternatively be written as $\mathbf{v} = \frac{\dot{a}}{a}\mathbf{r}$. This leads us to define the Hubble parameter (H) in Hubble's Law ($\mathbf{v} = H\mathbf{r}$) as,

$$H = \frac{\dot{a}}{a} \tag{1.9}$$

Thus, we see that the Hubble parameter is a direct measure of the expansion rate of the Universe and a vital parameter with which to constrain cosmological models. From the cosmological principle we will assume therefore that H is uniform over all space, however the expansion rate may (and must in all but a Steady State Universe) vary over time. We call the value of H today the Hubble constant (H_0).

The inverse of the Hubble constant ($1/H_0$) is known as the Hubble time (t_H), this would represent the age of the Universe (t_0) if the expansion rate had been constant. Current evidence suggests that the expansion of the Universe is accelerating, as such t_0 is expected to be less than t_H.

Due to the limited accuracy of the observations and overwhelming proper motions of the galaxies available to Hubble, measurements made in 1929 (Fig. 1.1) give a value of H_0

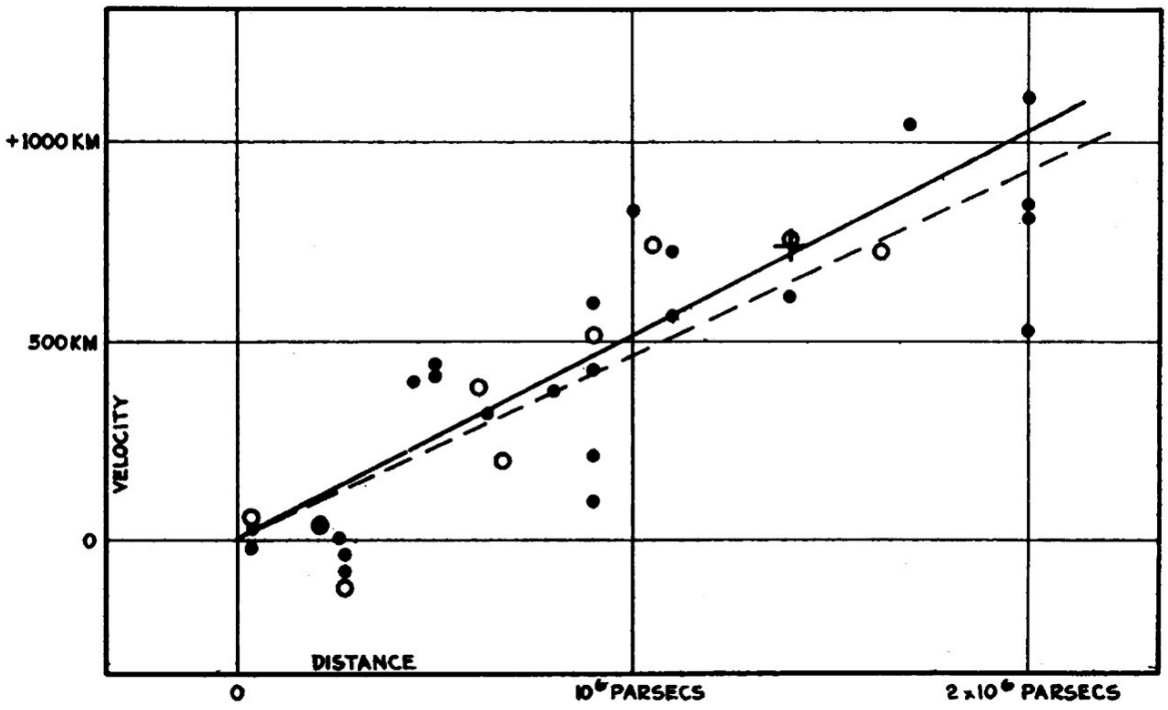


Fig. 1.1 Velocity-Distance relation for galaxies (then called "Extra-Galactic Nebulae") from the original Hubble (1929) paper. Full line from using galaxies individually (black dots), dashed trend from binned groups (open circles).

of around 530 $\mathrm{km\,s^{-1}\,Mpc^{-1}}$. More recent studies have constrained the Hubble constant at much lower values; $H_0 = 67.8 \pm 0.9\ \mathrm{km\,s^{-1}\,Mpc^{-1}}$ using the CMB (Planck Collaboration et al., 2016), $67.3 \pm 1.0\ \mathrm{km\,s^{-1}\,Mpc^{-1}}$ from measurements of the Baryon Acoustic Oscillation (BAO) signal in galaxy surveys (Alam et al., 2016), $73.24 \pm 1.74\ \mathrm{km\,s^{-1}\,Mpc^{-1}}$ using type 1a supernova (Riess et al., 2016), and $71.9^{+2.4}_{-3}\ \mathrm{km\,s^{-1}\,Mpc^{-1}}$ given by time delays from multiply imaged strong lensed quasar systems (Bonvin et al., 2017). Accounting for the disparity between these observations remains the subject of much ongoing research and controversy.

1.2.2 Cosmic Distance-Ladder Methods

Distances were first measured directly by astronomers via the parallax effect caused by Earth's relative change in position around the Sun. This method originally defined the parsec (pc) unit of measurement, with 1pc set as the distance at which a parallax of 1 arc-second was induced by the Earth's orbit. Due to the small angle approximation, the tangential relationship between parallax angle and distance in parsecs can be approximated as inversely proportional above $\sim$ 1pc. Hence, even using the highest resolution telescopes currently available, measurements of sources further away than a few hundred parsecs are unreliable and a new approach is required.

Beyond the limit of the parallax method there is no direct way to measure distance. Instead most methods rely on identifying standard candles, a group of sources which have some known brightness. A distance can then be determined using an object's luminosity and applying an inverse square law for a given cosmology. In general, objects in the Universe will have a distribution of brightnesses, even very similar stars will very with mass, age and composition. As such, correlations between other observables such as size and colour are used, along with physical assumptions to infer the intrinsic brightness. Objects within the galaxy for example, can be fitted to their evolutionary path on the Hertzsprung–Russell diagram (Webb, 1991).

Testing a luminosity relation on local sources, where it can be verified by a robust independent method, increases the number and variety of sources with a measurable distance. This then allows the discovery of new correlations which are valid over a different range of distances. For example, The period-luminosity relation of Cepheid variables (Leavitt, 1908) is one of the key techniques for measuring distance to nearby galaxies. There is a high correlation between the brightness and pulsation period of Cepheid variables of the same

type. This initial discovery, and the subsequent calibration of the Cepheid method, is relent on the shorter range measures discussed above.

There are a multitude of different relations and correlations for the various scales in the Universe. Collectively, these are known as cosmic distance ladder methods since each step is dependent on measures at smaller scales. As such, a measure of larger distances will inherit any uncertainty or bias from the methods on which is it based. This effect can be somewhat reduced by combining the results from various independent distance ladders.

The work in this thesis primarily uses redshift as a distance measure, as such I will now skip other methods. For more details on distance-ladder methods see Webb (1991), a discussion of Type 1a Supernovae is also included in section 1.4.3.

Redshifts

Using distance ladder methods on a wide range of scales, the velocity-distance relation of galaxies has been confirmed and constrained to much higher precision than presented first in Hubble (1929). Due to the Doppler effect this recession velocity causes the wavelength of a galaxy's spectrum to be increased by $\Delta\lambda = \frac{\lambda_e H r}{c}$, where r is the distance travelled by the photon. This lengthening of the wavelength induces a shift toward the red end of the (visible) spectrum, hence the effect is called red-shifting. The redshift distance (z) can then be used as a proxy for cosmological distances. This is defined in the standard terminology (Liddle, 2015) as,

$$z = \frac{\lambda_r - \lambda_e}{\lambda_e} = \frac{H}{c} r \tag{1.10}$$

Since the rest wavelength of emitted galaxy radiation (λ_e) is unknown, it must be estimated by identifying emission and absorption lines present in the galaxy's spectrum. These features are related to particular elements whose characteristic rest frequency can be identified through lab experiments. In practice, a catalogue of models representing different astrophysical components is used to compare spectra and identify the most likely redshift. For a more detailed description of spectral modelling and redshift identification the reader is directed to Dawson et al. (2013).

As with all astronomical measurements, redshifts are subject to errors both in the data and post-observation analysis. The difference in wavelength, and hence the inferred distance, is not only dependent on a galaxy's recession velocity due to the expansion rate. There is

also a contribution from the peculiar velocity of the galaxy and (to a much smaller degree) effects along the line of sight e.g. the Sachs-Wolfe effect. However, non-expansion related red-shifting should not increase at the same rate as the expansion term. Thus, as distance increases the contribution from the Hubble flow dominates.

We can also consider a measure of the scale factor at the time of emission a_e as,

$$a_e = \frac{1}{1+z} \tag{1.11}$$

1.2.3 Distances in cosmology

The usual concept of distance, the idea that a straight line traces the shortest path between two points, assumes a Euclidean (flat and static) geometry. This simple definition does not properly describe distance on the surface of the Earth, let alone the curved and expanding space-time of the Universe. As such, there are several ways a distance can be defined for cosmic scales. Here, we describe four of the most common distances: proper, comoving, luminosity, and angular diameter.

Proper Distance

This distance relates the separation of an object emitting a photon at time t_1 to an observer at t_2; observationally redshifts can be used to express these times and so we define distances as a function of $z_1 = z(a(t_1))$ and $z_2 = z(a(t_2))$. Proper distance (D_p) is the quantity most closely related to our natural concept of a distance measure. It is given by the time taken for a photon to travel from z_1 to z_2 (multiplied by c). A definition for proper distance can be found by considering only the time-like part of the FRW metric, seen in eq. 1.3, such that,

$$dD_p = -cdt = cda\left[\frac{1}{aH_a}\right] \tag{1.12}$$

Integrating this between z_1 and z_2 gives,

$$D_p(z_1, z_2) = c\int_{z_1}^{z_2} \frac{1}{H_z(1+z)} dz \tag{1.13}$$

Comoving Distance

The proper distance is not often used in cosmology since it is dependent on time and evolves with the expansion rate. As such, a measure independent of expansion, called the comoving distance (D_c) is used. This is found by considering a null geodesic in FRW space-time, setting $ds = 0$ in eq 1.3, in effect removing a factor of a from dD_p. Thus,

$$dD_c = -ca^{-1}dt = cda\left[\frac{1}{a^2 H_a}\right] \tag{1.14}$$

such that,

$$D_c(z_1, z_2) = c\int_{z_1}^{z_2} \frac{1}{H_z} dz \tag{1.15}$$

Luminosity Distance

In a Euclidean space, or over short enough distances, the intensity of light can be seen to decrease proportionally to the distance travelled i.e $I \propto r^{-2}$. The Luminosity distance (D_l) is defined by assuming this holds for the Universe in general, and thus the flux (f) received by an observer at distance D_l from a source with luminosity L would be,

$$f = \frac{L}{4\pi} D_l^{-2} \tag{1.16}$$

$$\Rightarrow D_l = \frac{L^2}{(4\pi f)^2} \tag{1.17}$$

As a function of comoving distance however, the received flux would be less due to the effect of expansion not included in this measure. The surface area of a sphere in an expanding space would increase by a factor of a, decreasing the incidence of photons; the stretching of space time would reduce their energy by a further factor of a. Hence flux measured as a function of comoving distance would be proportional to $(ar)^{-2}$ giving,

$$f = \frac{L}{4\pi}(aD_c)^{-2} \tag{1.18}$$

$$\Rightarrow D_c = a^{-1}\frac{L^2}{(4\pi f)^2} \tag{1.19}$$

Thus, comoving and luminosity distances are related by,

$$D_l = a^{-1} D_c = (1+z) D_c \tag{1.20}$$

Angular Diameter Distance

Another observable which is dependent on distance is the angle subtended by an object on the sky. As with the luminosity distance we consider a static Euclidean space-time to make an initial definition. An object with physical extent l with angular size θ would have an angular diameter distance (D_A) of,

$$D_A = \frac{l}{\sin\theta} \simeq \frac{l}{\theta} \tag{1.21}$$

for a small value of θ which is generally the case in cosmological observations. In an expanding Universe the apparent angular size of an object will change in proportion with its size relative to a circle at that distance i.e by a factor of a. As such we find, $D_c = c/a\theta$ and all these distance measures are related by,

$$D_A = a D_c = a^2 D_l \tag{1.22}$$

Redshift is often used by astronomers as a simple proxy for distance. However, it is important to remember that there are other valid measures available and their various relationship to the underlying cosmology.

1.2.4 Observational Parameters

In order to constrain or rule out a particular model of the Universe, it is common to express a model using a few parameters. This is usually related to the expansion rate or composition of the Universe that can be measured observationally. This allows us to determine which model best fits the available data, or make conclusions about our Universe assuming one model in particular (usually ΛCDM).

Hubble constant H_0

We have already seen in section 1.2.1 that the Universe is expanding, and that this expansion has not been constant over time. We describe the time dependent expansion with the Hubble Parameter (H) and the value today as H_0 the Hubble Constant. The Hubble Parameter tells us the expansion history and scale of the Universe; it is therefore vitally important.

H_0 should be straight forward to measure since it is simply the ratio of distance and recession velocity of galaxies. However, as we have seen, measuring distance in a non-Euclidean space is not trivial, and reliable redshifts can be difficult to obtain for faint, distant galaxies and may be biased by other effects. The exact value of H_0 is not agreed upon to high precision; it (and other dependent quantities) is often given as a multiple of h where $H_0 = 100h\mathrm{km\,s^{-1}\,Mpc^{-1}}$. The Hubble Constant is measured in inverse time units of $\mathrm{km/\,sec\,/Mpc} \equiv 3.2 \times 10^{-20}\mathrm{sec^{-1}}$ related to the age of the Universe.

There is some tension in the values of H_0 between probes of the local (SNe) and early (CMB) Universe. This can be seen in table 1.1. Despite this, measurements give a broadly consistent value constraining one of the fundamental parameters of cosmology to within a few percent.

Density Ω

The density parameter is a useful way of analysing the evolution of density over time and its relation to the curvature of the Universe. Considering Friedmann's equation (1.5) again we can express the density of the Universe as,

$$\rho = \frac{1}{8\pi G}\left(3H^2 - \Lambda + 3Kc^2a^{-2}\right) \tag{1.23}$$

A universe with no cosmological constant and zero curvature therefore has a specific density. This is called the critical density ρ_c.

$$\rho_c = \frac{3H^2}{8\pi G} \tag{1.24}$$

Since the Universe is expanding ρ_c will also be a function of time, but we can use the current value of H to find what the critical density should be today.

$$\rho_c(t_0) = 1.88h^2 10^{-26}\mathrm{kg\,m^{-3}} \tag{1.25}$$

This provides a reference by which to compare the density of our Universe. The density parameter is thus defined as the ratio of the measured density and the critical density, $\Omega = \rho/\rho_c$. Again we use the notation Ω_0 to denote the present value of Ω.

The density parameter is also a useful way of describing the ratio of the other components of the Universe which we will cover in the next section. Defining an energy density parameter for curvature,

$$\Omega_k = -\frac{k}{a^2H^2} \tag{1.26}$$

which can be positive or negative, results in,

$$\Omega + \Omega_k = 1 \tag{1.27}$$

Density dispersion σ_8

As the Universe has evolved, gravitational collapse has transformed the density field from uniform with small ($\sim 10^{-5}$) perturbations, to one with great contrasts. We define this fluctuation as σ_8, the standard deviation of the over-density field smoothed on $8h^{-1}$Mpc scales.

Since the formation of structure depends on how matter interacts, the energy densities, and the evolution of the Universe, σ_8 is sensitive to many fundamental properties. As such it is very useful for comparing the results of cosmological models to observations.

1.3 The ΛCDM Universe

This section gives a broad overview of the current consensus model of cosmology, ΛCDM. I describe the main physical components of the model, recognising the gaps in our knowledge and the best theories explaining them. I also look at some alternatives which could be tested.

1.3.1 Energy components

As shown in section 1.2.4, Ω gives the density of all energy (apart from curvature) in the Universe. This density can be broadly split into relativistic energy, matter, dark energy, plus a scalar field required for inflation. As seen in equation 1.27 the sum of all the energy density values only differs from 1 by a quantity dependent on curvature. Hence, in a near flat universe such as ours,

$$\Omega = \Omega_r + \Omega_m + \Omega_\Lambda \simeq 1 \tag{1.28}$$

As the Universe expands the densities are all diluted at a rate dependent on their equation of state (w) as,

$$\Omega_i(t) = a^{(-3(1+w_i))}\Omega_i^{(0)} \tag{1.29}$$

Relativistic energy

The relativistic portion of the energy density comes from electromagnetism radiation ρ_γ and neutrinos which behave relativistically when their masses are low enough (ρ_ν). This relativistic energy can be calculated using the evolution of pressure from equation 1.7. We generally refer to this type of energy as radiation using the density parameter Ω_r. The pressure and density of radiation is related by $p = \rho/3$ giving an equation of state of $w = 1/3$. From equation 1.29 we can see that Ω_r dilutes very quickly with time as,

$$\Omega_r = a^{-4}\Omega_r^{(0)} \tag{1.30}$$

The factor of a^{-4} comes from the dilution in space as the Universe expands by a^3 coupled with a loss in energy due to redshifting (see section 1.2.2). Although radiation dominated in the early Universe, the rapid dilution means that the contribution of CMB photons today is $\Omega_\Lambda^{(0)} \simeq 5 \times 10^{-5}$ and is often neglected.

The density of relativistic particles is given by Amendola & Tsujikawa (2010) as,

$$\rho_\nu = N_{eff}\frac{7\pi^2}{120}T_\nu^4 \tag{1.31}$$

where N_{eff} is the number of effective neutrino species and T_ν is their temperature. The standard model of particle physics predicts three species of neutrino (electron ν_e, muon ν_μ, and tau ν_τ). In ΛCDM models N_{eff} is generally taken to be ~ 3.04 (Mangano et al., 2002). Some models do allow N_{eff} to be a free parameter introducing a so called "dark radiation" as a possible solution for the tension seen in measures of H_0.

Non-relativistic Matter

Non-relativistic matter was thought for a long time to be composed of only baryonic particles in the form of gas, dust, stars and other things that interact with light. However, work by Zwicky (1937) to estimate the masses of galaxy clusters found an inconstancy between a value based on luminosity and a value derived from the distribution and movement of internal galaxies. In particular, studies of the Coma cluster provided the first evidence for the

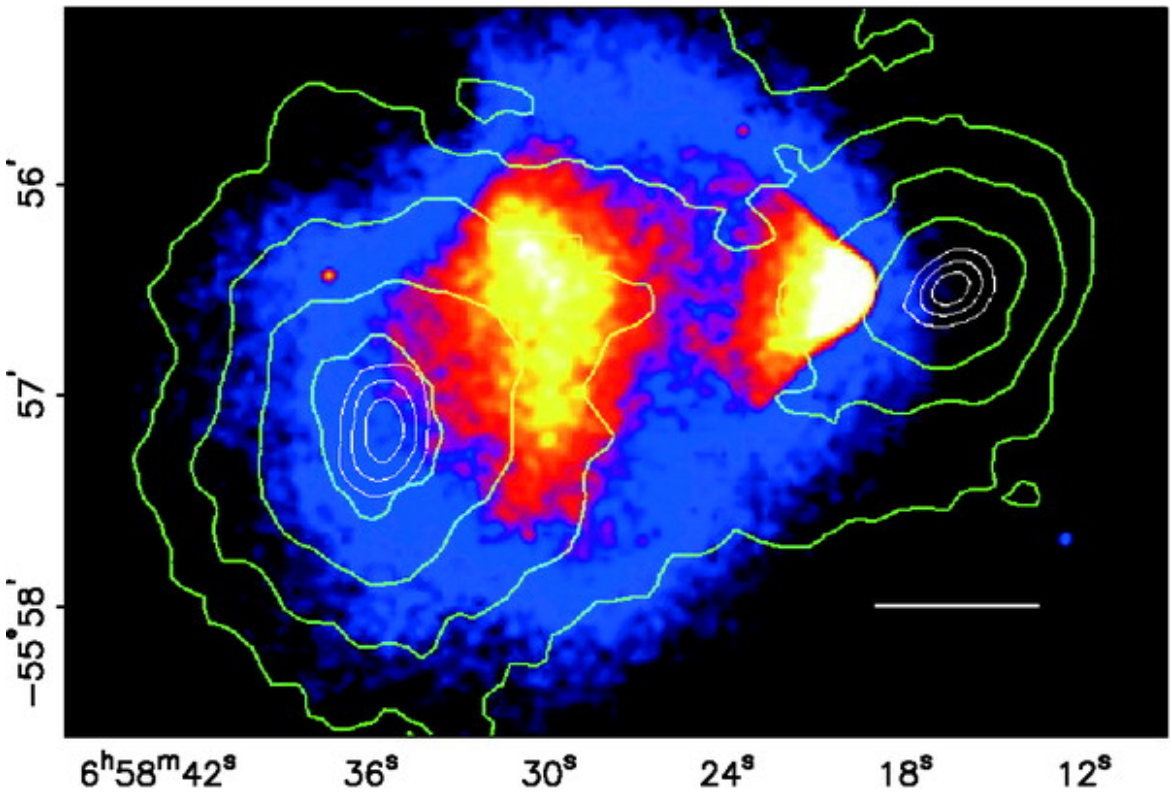


Fig. 1.2 X-ray image from Chandra of the Bullet Cluster with green contours overlaid from the weak-lensing κ reconstructions. Outer contour levels at $\kappa = 0.16$ and increasing in steps of 0.07. Figure taken from Clowe et al. (2006).

existence of a type of matter which interacts with gravity, but not light "dark matter".

This work was later confirmed by Rubin & Ford (1970) who provided much stronger evidence for the presence of dark matter through the study of galaxy rotation velocities. Ruben's work demonstrated that dark matter must constitute a significant proportion of the mass within galaxies, and be essential for their formation and stability. Thus, dark matter was included in the standard cosmological model with Ω_m separated into two parameters (Ω_b and Ω_{DM}) representing baryonic and dark matter respectively.

Dark matter has yet to be detected directly, but further evidence for its existence has been discovered though the use of weak lensing (see Chapter 4). The most famous example of this is from the observations of cluster 1E 0657-558 (the Bullet Cluster) by Clowe et al. (2006) seen in Fig. 1.2. The collision of two large clusters separated the dark matter and baryons, leading to a discordant profile between gravitational and thermal interaction.

We now consider the various candidates which have been suggested for dark matter.

Observed Candidates:

- *Massive Astrophysical Compact Halo Objects (MACHOs)*: One suggestion was that dark matter consists of normal atomic matter "hidden" in very dense regions called MACHOs that emit little or no light e.g. black holes or brown dwarf stars. However, evidence from primordial nucleosynthesis (Larson et al., 2011) gives a density for baryonic matter of only $\Omega_b \simeq 0.045$. This is much lower than the total matter density of $\Omega_m \simeq 0.3$ measured using the CMB, BAO and SNe 1a, see table 1.1. Work by Alcock et al. (2000) has further constrained the contribution of MACHOs to dark matter through the detection of microlensing events. They find that, for a typical model, MACHOs account for 20% of the dark matter in the Milky Way's halo. A 95% confidence interval of $8-50\%$ rules out a Galactic halo comprised entirely of MACHOs.

- *Massive neutrinos*: Massive neutrinos only interact through gravity and the weak nuclear force. Neutrinos have been directly observed, making them an obvious candidate for dark matter. However, the upper limits on neutrino masses mean that they have velocities of almost the speed of light (ultra-relativistic), hence they are called Hot Dark Matter (HDM). Simulations of the Universe using HDM do not form the structures that we see from observation, on the correct time scales. This is due to the particles streaming out from the smaller over-densities, preventing gravitational collapse. As such HDM is unlikely to be a significant component of dark matter.

Theoretical particles:

- *WIMPs (CDM)*: The most commonly accepted form for dark matter is weakly interacting massive particles (WIMP)s that unlike neutrinos are non-relativistic in the early Universe (Arkani-Hamed et al., 2009). We refer to this type of WIMP as Cold Dark Matter (CDM). CDM has been shown to collapse in simulations in a way which is consistent with observations of present structure. Although it has not been detected, CDM is the best fit to current data and as such is currently the favoured form of dominating matter.

- *Supersymmetry*: It has been suggested that all ordinary atomic particles have a counterpart at much higher mass and energy (Liddle, 2015). In many theories the lightest of these supersymmetric particles would be a zero charge WIMP, and may be detected by the Large Hadron Colider (LHC) in the near future.

- *Axions:* Finally, axions are another theoretical particle that interact weakly due to their very small mass and couple to photons in the presence of magnetic fields. Axions would have an energy in between WIMPs and massive neutrinos and are hence a possible candidate for Warm Dark Matter (WDM), (Covi et al., 2001).

These types of non-relativistic matter have a negligible pressure giving an equation of state of $w = 0$, hence the energy density of matter falls with the volume of the Universe.

$$\Omega_m = \Omega_b + \Omega_{DM} = a^{-3}\Omega_m^{(0)} \tag{1.32}$$

Dark Energy

In ΛCDM the remaining energy density and the "force" driving accelerated expansion in the Universe is called dark energy, which may be a cosmological constant Λ. The constraints placed on Ω_m and Ω_k combined with equation 1.27 mean that this component must have an energy density of $\Omega_\Lambda \simeq 0.7$ today. This means around 70 percent of cosmic energy is unidentified in the standard model. In the next section we will consider some of the alternatives to Λ for dark energy.

Assuming Λ is a constant, then the fluid equation (1.7) gives $p = -\rho$ and an equation of state of $w = -1$, hence the density parameter of dark energy in this model is,

$$\Omega_\Lambda = \Omega_\Lambda^{(0)} \tag{1.33}$$

and dark energy does not evolve with time. This means that although Λ dominates in the recent Universe, at early times it would have been a much smaller component than matter or radiation. We can consider a more general version of ΛCDM which features an evolving dark energy component by not assuming $w = -1$ and fitting the observed evolution of dark energy to,

$$\Omega_\Lambda = a^{-3(1+w_0)}\Omega_\Lambda^{(0)} \tag{1.34}$$

We call these models wΛCDM. Observations from the CMB combined with SNe Type 1a have measured the equation of state as $w = -1.006 \pm 0.045$ (Planck Collaboration et al., 2016) consistent with a constant Λ.

1.3.2 Dark Energy Effect without a Cosmological Constant

There are many alternative theories to explain cosmic acceleration that don't include an arbitrary constant in the fundamental equations of gravity. These explanations range from an additional particle field such as "quintessence", to multi-dimensional braneworld models, to abandoning the cosmological principle of homogeneity for certain scales altogether. Here we take a broad look at three areas that could explain the nature of dark energy. For more details on these theories, or Dark Energy in general, I refer the reader to Amendola & Tsujikawa (2010).

Vacuum Energy

It has been long observed from particle physics laboratory experiments that empty space produces a negative pressure due to quantum fluctuations, a so called vacuum energy. This quantum field would produce the accelerated expansion attributed to dark energy, but at a density $\sim 10^{121}$ times the observed value. Rather than solving the problem of cosmic acceleration, vacuum energy introduces a second, much larger, one. However, if further understanding of quantum interactions on large scales reduces the expected effect, or some reasonable damping mechanism can be found, this could be a solution for Λ.

Modified Gravity

As earlier discussed, some models consider a more general form of Λ by not fixing its equation of state. This can be extended further by modifying the Einstein tensor with some general function. The aim is to replicate late time expansion without an explicit dark energy component. For the most part modifications that still satisfy solar system constraints do not produce pronounced effects and are therefore difficult to test observationally. However, tests are currently under-way, for example, Burrage & Sakstein (2017).

Cosmic acceleration from voids

Finally, some cosmologists suggest that universal dark energy does not really exist at all and is in fact an observational effect. This clearly breaks the cosmological principle but would explain the variable and recent nature of dark energy. Also, there is no reason why a strong local inhomogeneity could not still preserve isotropy. One way to simulate dark

energy using an inhomogeneity is by way of a large spherical void, or many stacked voids, in which our galaxy is at the centre. Garcia-Bellido & Haugbølle (2008) find that current observational evidence cannot rule out such a structure, although it is difficult to account for the entire effect of Λ in this way. There is currently no valid mechanism for the formation of inhomogeneous structure on this scale, and models often rely on coincidence and degeneracy. However, theories like void stacking that explain acceleration with no dark energy component do pose an important question; is the smaller picture sometimes more important than the big one?

1.4 Probes of Cosmology

A cosmological probe is any measure of the Universe that can be used to constrain the cosmological parameters defining the ΛCDM model, or test a prediction of an alternative model. The constraining power of a single probe is often weak when applied to many parameters simultaneously. As such, a smaller number of them are sometimes fitted, making a reasonable assumption for the remainder. For example, it is possible to fix Ω_k and constrain the parameter space of $w - w_0$. By combining multiple probes that are sensitive to different parameters, we are able to form a complete picture of cosmology. We now discuss some of the probes that have been important in developing the concordance model, and a few that may be significant to the cosmological model of the future.

We summarise the latest constraints on cosmology from these various probes at the end of the section in table 1.1.

1.4.1 Cosmic microwave radiation

In any direction observed from the Earth a small, almost perfectly uniform, signal can be seen, the most recent measurement of which can be seen in Fig. 1.3. This microwave frequency radiation is the oldest and furthest light detectable in the Universe and is known as the Cosmic Microwave Background (CMB). The CMB is a crucial piece of evidence supporting not only the hot Big Bang theory as the origin of the Universe, but also the cosmological principle as the foundation of cosmology.

Alpher et al. (1948) first theorised that the early Universe, in a Big Bang model, would have to be so hot that photons and baryons would be strongly coupled as a plasma thorough

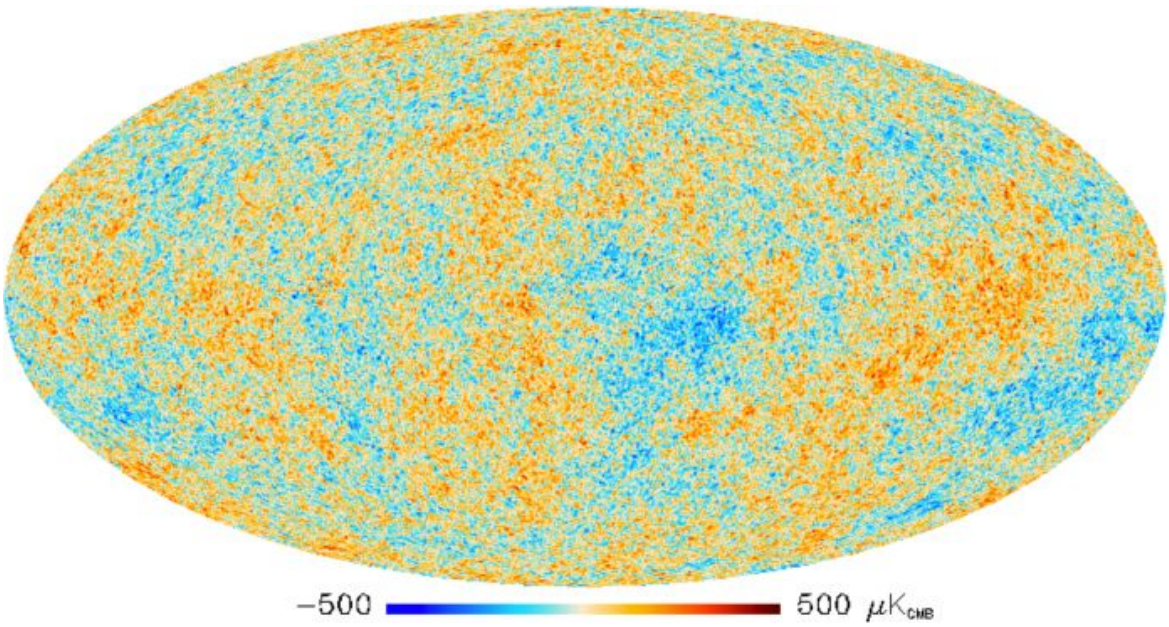


Fig. 1.3 CMB temperature fluctuations relative to ~ 2.7 Kelvin as measured by Planck Collaboration et al. (2016)

which light could only travel a short distance. As the Universe expanded and cooled to around 3000K photons would not have enough energy to prevent atoms of neutral hydrogen from forming. Over a relatively short period of time, the Universe turned from opaque to transparent as light, unimpeded by ionised plasma, was able to free-stream to us in a process known as decoupling. These photons all originate from the same time, and hence distance, the inside of a sphere called the last scattering surface or CMB. The CMB has been further cooled and red-shifted to the mm wavelength, 2.7K signal we observe today.

The CMB has been measured several times since its initial, Nobel prize winning, discovery by Penzias & Wilson (1965). The Cosmic Background Explorer (COBE; Smoot et al. 1991) was the first space based instrument designed to measure anisotropies in the CMB. Its discovery that the temperature distribution of the CMB is almost uniform, with deviations of around 10^{-5}, imply that the Universe was homogeneous, at the time of decoupling, over distances too large to be causally connected. This so called horizon problem led Guth (1981) to suggest that the uniformity seen by COBE could be caused by "Inflation", an expansion of the very early Universe by a factor of around 10^{30}.

Anisotropies in the CMB were measured more accurately by the Wilkinson Microwave Anisotropy Probe (WMAP; Komatsu et al. 2011) and then most recently by the *Planck*

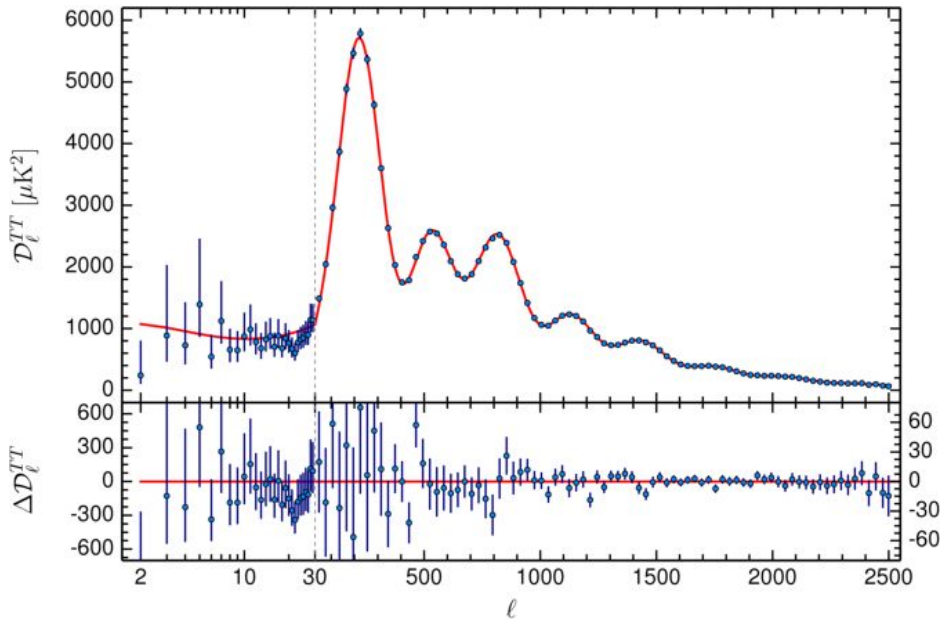


Fig. 1.4 Planck Collaboration et al. (2016) CMB temperature power spectrum with 1σ error bars. Red line shows best fit ΛCDM model spectrum to the data with residuals shown below.

experiment (Planck Collaboration et al., 2016). One method for analysing the CMB is to measure the power spectrum, a two point statistic of the temperature fluctuations as a function of angular distance ℓ. The power spectrum calculated from the most recent observations by the Planck Collaboration et al. (2016) is reproduced in Fig. 1.4. By fitting a power spectrum model based on a ΛCDM Universe to the data we are able to constrain the energy density parameters and the Hubble parameter.

The position and relative heights of the small scale (ℓ>500) peaks seen in the power spectrum are subject to the ratio of energy densities and, at $\ell > 1000$, small scale physics through which the peaks are damped by photon diffusion. Higher order measurements of the temperature field can be used to probe a combination of h with Ω_r, Ω_m, Ω_b, and Ω_Λ as described by Planck Collaboration et al. (2014).

The position of the first peak at $\ell \simeq 200$ however, is directly related to the scale of the last scattering surface i.e the inverse of the Hubble time H^{-1} at that redshift. We can calculate the redshift of the CMB since we know its rest spectrum was that of a black-body at a temperature that would allow decoupling to occur. Hence, we have an explicit indication of the scale and thus curvature energy (Ω_k) and geometry of the Universe at the time of last scattering. The results of the Planck Collaboration et al. (2016) observations show that the curvature of the Universe is well within a percent of zero with $|\Omega_k| < 0.005$ explicitly ruling

out significantly non-flat geometries.

Photons from the CMB rarely interact with intervening matter along the line of sight, and are as such mostly sensitive to physics at z$\simeq$ 1100. Despite this the weak interaction of the CMB through the Sunyaev-Zeldovich effect (Sunyaev & Zeldovich, 1970) and the Integrated Sachs-Wolfe effect (Sachs & Wolfe, 1967) provide a probe of the late Universe. Other contaminations such as foreground emission from the local Universe and the dipole due to the relative motion of the Solar System can be estimated and removed through multi-wavelength data.

1.4.2 Baryon Acoustic Oscillations (BAO)

The CMB also has implications for structure in the present day Universe. Before recombination the Universe was a dense plasma and pressure waves were able to propagate. This created over-densities of matter in spherical shells. When photons decoupled from atoms these waves could no longer travel becoming "frozen" in the matter distribution of the early Universe. These small over-densities were the seeds of structure formation, with stars and galaxies more likely to form around them due to gravitational collapse. The scales on which this structure forms is dependent on the energy density make-up of the Universe at the time (Bond & Efstathiou, 1984; Holtzman, 1989; Peebles & Yu, 1970; Silk, 1968).

Today this signal can be measured by way of a two-point statistic of galaxies as a function of their co-moving distance. In order to detect this BAO signal a high quality wide field spectroscopic survey of 10^5 to 10^6 galaxies was needed. The SDSS Baryon Oscillation Spectroscopic Survey (BOSS) was designed to meet this requirement. The measurement of the BAO by Ross et al. (2017) is reproduced in Fig. 1.5, where a peak can clearly be seen at a separation of $\sim 100h^{-1}$Mpc. This means that any two galaxies are more likely to separated by $\sim 100h^{-1}$Mpc than by slightly less or more.

The BAO peak can also become damped and distorted on large scales by the non-linear peculiar motions of galaxies (Angulo et al., 2008). In order to overcome this increased uncertainty the linear BAO is reconstructed by removing as much bulk motion as possible; for an analysis of this technique see Burden et al. (2014).

It has also been proposed that the BAO could be measured in the radio by mapping neutral hydrogen gas, as traced by the 21cm line. Baryon acoustic oscillations In Neutral Gas

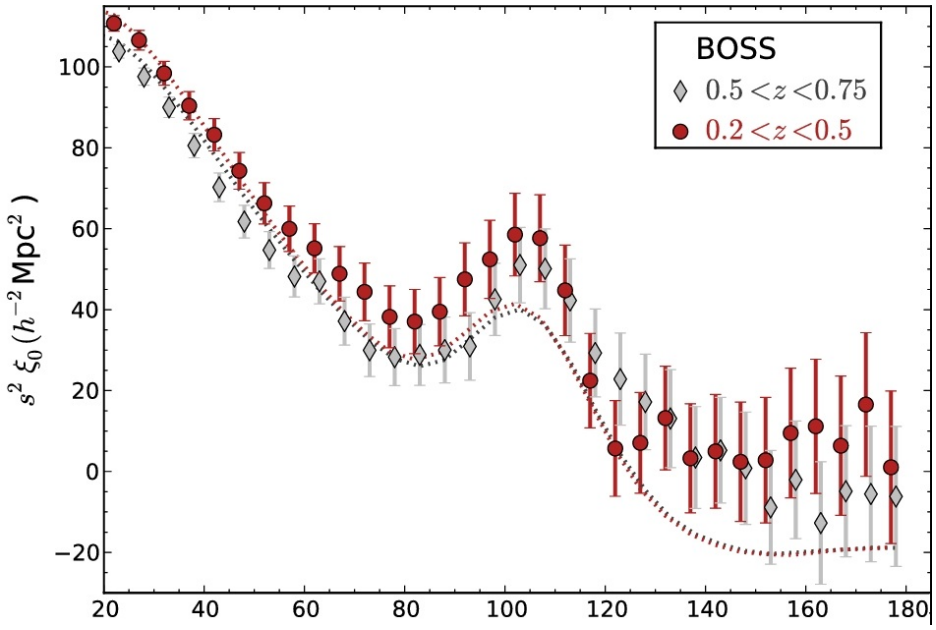


Fig. 1.5 The galaxy correlation monopole as a function of co-moving separation (h^{-1}Mpc) measured from BOSS DR12 in two redshift bins. Figure taken from Ross et al. (2017). Red dashed line shows the MultiDark PATCHY mock realisation of the BAO by Kitaura et al. (2016, 2014).

Observations (BINGO) is a planned intensity mapping survey that Dickinson (2014) suggest could measure the acoustic scale to ~ 2 percent precision. Although this is less accurate than optical surveys have already achieved, the independent nature of the measurement would be complementary to existing data reducing systematic bias from probing galaxies alone.

1.4.3 Type 1a Supernovae

There is no measure for distances along the line of sight, of over $\sim$kpc, that is independent of physical processes. As such, the primary source of noise for a measurement on the Hubble diagram is the distance estimation. If the galaxy brightness distribution was constant with redshift, then the Euclidean normalised source count as a function of magnitude could be used as a statistical proxy for the proper distance. However, there is too much variability in the general galaxy populations for this to be feasible.

Type 1a Supernovae are all caused by a very similar physical process, that of either a white dwarf star accreting gas from a companion, or the merger of two white dwarfs. A supernova explosion occurs when the total mass of the stars involved reaches the Chandrasekhar limit

of $\sim 1.4 M\odot$ (Chandrasekhar, 1931). The characteristic intrinsic luminosity of a Type 1a shows some dispersion in the peak brightness.

However, it was shown these discrepancies could be standardised through the relationship to the rate at which the light curve decays (Phillips, 1993), and the colour of the supernovae (Riess et al., 1996; Tripp, 1997). In effect this creates "standard candles" which can be used to measure the distance modulus, a quantity constructed from the observed magnitude and accurately related to the luminosity distance. Type 1a Supernova also have high intrinsic luminosities of over 10^9 times that of the sun often making them brighter than their galactic host, making it possible to observe them out to very high redshifts.

Despite being referred to as standard candles, Type 1a are not perfect rulers for the Universe. Firstly, the standardisation of a supernova requires multiple observations over the lifetime of the explosion, preferably beginning before reaching peak brightness. We never know exactly when a supernova will occur, so continuous scanning surveys combined with rapid follow-up is necessary. Secondly, a redshift is still needed in order to construct a distance modulus diagram and probe the parameters on which it depends. Redshifts are usually obtained through high quality spectroscopic observations of the host galaxy which are difficult at high redshift, time consuming and expensive. Using photometric redshifts for the bulk of nearby classifications is a possible alternative, with Campbell et al. (2013) demonstrating competitive statistical errors. Finally, work by Howell et al. (2006) showed that some "overluminous" Type 1a supernovae cannot be calibrated by the accordant relation used for standardisation.

The distance modulus diagram constructed from a collection of supernovae observations complied by Suzuki et al. (2012) can be seen in Fig. 1.6. From this data multiple teams (Garnavich et al., 1998; Perlmutter et al., 1998; Riess et al., 1998; Schmidt et al., 1998) have concluded that the Universe is undergoing accelerated expansion. A model of this diagram assuming a matter only universe is a poor fit to the data, underestimating the distance modulus particularly at high redshift. This implies that the cosmological constant Λ is non zero, with a best current estimate of $\Omega_\Lambda = 0.729 \pm 0.014$.

1.4.4 Time delay strong lensing (TDSL)

Gravitation lensing is caused by the bending of light paths around massive objects on the line of sight. Lensing causes images of distant galaxies (and the CMB) to change shape and

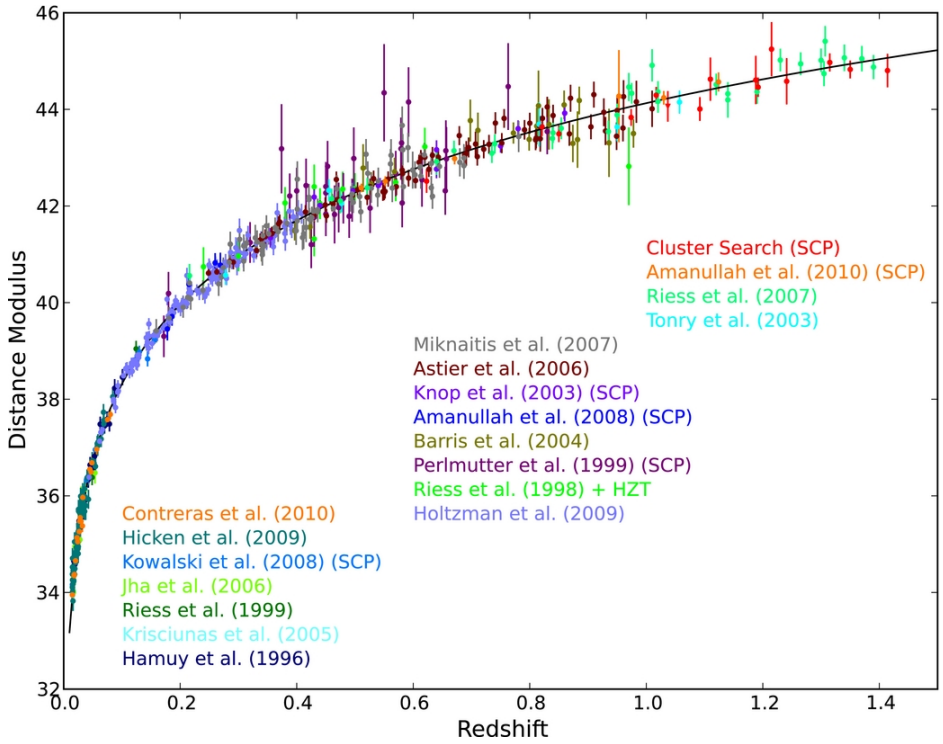


Fig. 1.6 Distance modulus diagram from compilation of several supernovae surveys as seen in Suzuki et al. (2012).

size. We discuss the background of lensing in more detail in chapter 4.

When a background galaxy is in the right position behind a massive object, like a galaxy cluster, its image can form large arcs, Einstein rings, or multiple images as seen in Fig. 1.7. We call this type of distortion strong lensing. When a strong lens produces multiple images of the source, the paths taken by the lensed light are different lengths, inducing a time delay between the various images. If the multiply imaged source varies over time e.g. a supernovae or a quasar with non-periodic variability, it can therefore be used to measure the time delay effect, a quantity which is dependent on the cosmological properties of the intervening Universe (Bonvin et al., 2017).

Since TDSL is a direct probe of the geometry of the Universe Suyu et al. (2013) find that by using only two strong lenses they are able to constrain Ω_Λ, Ω_k, and w to levels comparable to BAO. In particular, they find TDSL is sensitive to the Hubble constant, able to constrain H_0 to 5.7 percent for wCDM and 4 percent assuming $w = -1$ and $\Omega_k \geq 0$. More recently Bonvin et al. (2017) constrained H_0 even further to ~ 3.5 percent for ΛCDM, also showing that TDSL can help break degeneracies faced by *Planck* and BAO in constraining more

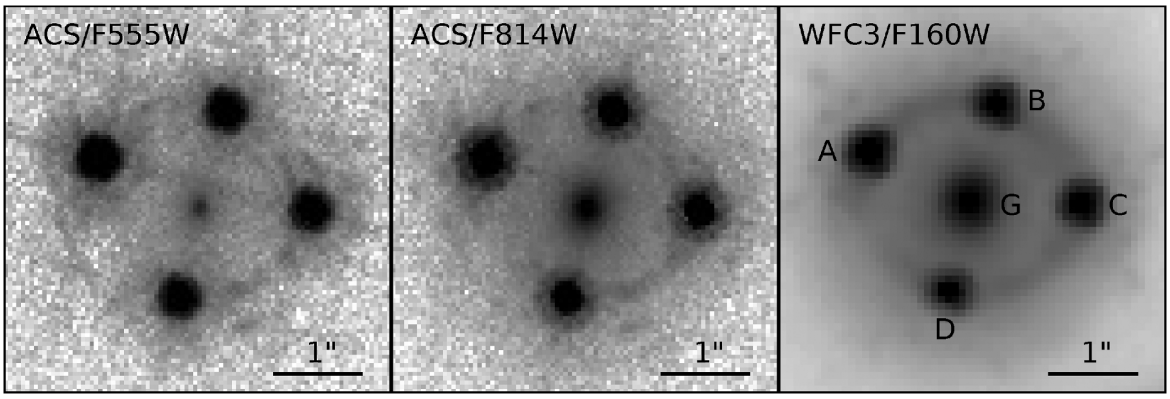


Fig. 1.7 Cutout from the Hubble Space Telescope (HST) of HE 0435-1223 a lens galaxy (G) and a background quasar source lensed into four images (A to D). Figure taken from Wong et al. (2017).

general models such as wCDM, NeffΛCDM and m$\nu\Lambda$CDM.

The main obstacle facing improvements in TDSL measurement is simply the lack of strongly lensed sources suitable for time delay analysis. There are only a few hundred high quality strong lenses currently known, and although Collett (2015) predicts this to rise to $10^{3.3}$, 10^5 and $10^{5.3}$ in DES, LSST and Euclid respectively, identifying them will still be a challenge as they only account for around 10^{-6} of sources. Classifying lenses by way of convolutional neural networks has been suggested as one solution by Jacobs et al. (2017), however this technique currently only reaches completeness of ~ 25 percent.

1.4.5 Weak Lensing

The effect of lensing near smaller gravitational potentials can also be a powerful probe of cosmology. In this regime the distortion is not strong enough to produce multiple images or large arcs from sources. The main effect of weak lensing is a slight change to galaxy ellipticities. This cannot be measured directly but can be used to make a statistical inference of cosmological parameters. Weak lensing will form a large part of the work in the second half of this thesis, and as such I defer the details of this measure to chapter 4.

1.4.6 Gravitational waves

One of the key predictions of General Relativity is that gravitational interactions will propagate as a wave. Evidence for this was first discovered indirectly by Taylor et al. (1974)

though observations of a binary pulsar. More recently a gravitational wave was directly observed for the first time by the Laser Interferometer Gravitational-Wave Observatory (LIGO) (Abramovici et al., 1992) seen in Fig. 1.8. This detection not only helps confirm the theory of General Relativity, but also extends our ability to probe the Universe beyond the electromagnetic spectrum.

Gravitational wave observations are still in their infancy and have not yet had time to contribute to constraints of cosmology. However, even at this early stage the high rates of binary black hole mergers point to ignorance in our model of formation history. In the future, measurements of the stochastic gravitational wave background along with follow-up of individual events by traditional telescopes, will provide the most powerful constraints for studies of the early Universe and cosmic inflation, (Maggiore, 2000).

Above, I have surveyed key cosmological probes which give evidence of dark matter and dark energy. Current constraints on cosmological parameters are found in table 1.1. While this progress is impressive, the next generation of surveys will be able to measure these parameters with much more fidelity. In this thesis I seek to contribute to this progress.

1.5 Overview of this Thesis

In the next chapter I provide details about radio astronomy that will be essential for the later work in this thesis. I begin by describing the various physical emission processes that produce the observable EM field. I then consider radio instrumentation and the methods of collecting and interpreting data. I show the relationship of the observed sky brightness distribution to the radio "visibilities" and discuss the traditional method of inverting it. Finally, I give an overview of radio galaxy population studies as a precursor to the work in chapter 3.

In chapter 3 I present the results of a spectroscopic observation of radio selected sources. I led the team that originally conducted this work, and the chapter is based on the publication it has produced, (Tarr et al. 2017a). This study formed a pilot for the WEAVE-LOFAR survey commencing in 2018 (Smith et al., 2016).

Chapter 4 gives an overview of gravitation lensing. In particular, I discuss the theoretical basis of weak lensing, building to the distortions that are expected from gravitational interactions in this regime. I consider radio observations as a future alternative to traditional optical

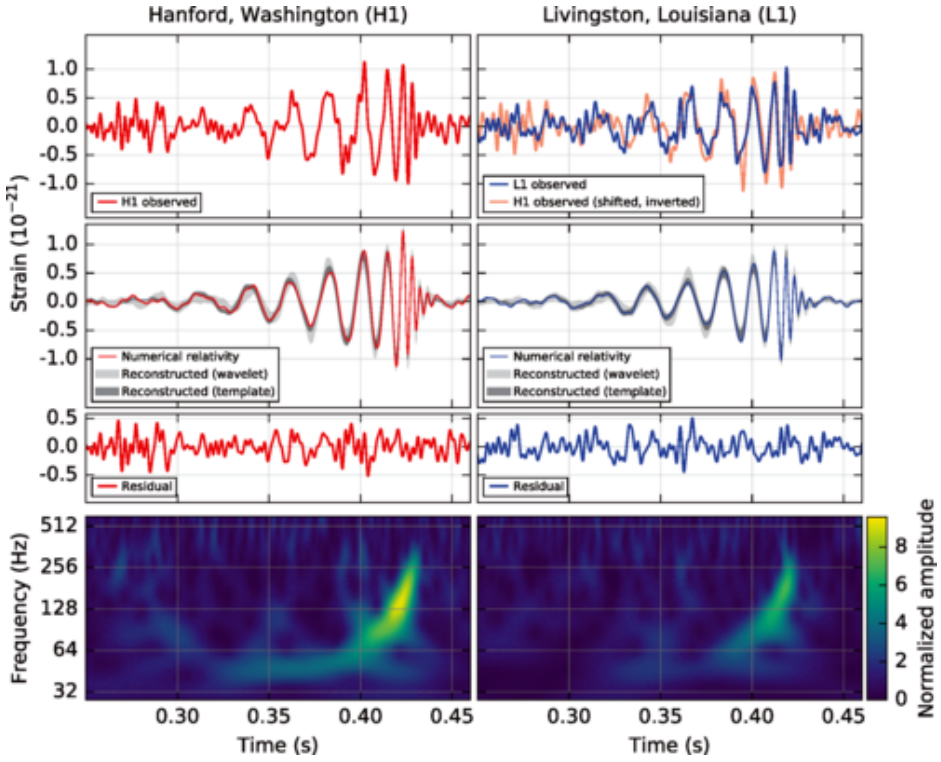


Fig. 1.8 First directly detection of a gravitational-wave event (GW150914) caused by a binary black hole merger. Observed by the LIGO detectors September 2014. Figure taken from Abbott et al. (2016).

lensing studies, and discuss the advantages and challenges this presents. I conclude with the current predictions for SKA-cosmology from radio weak lensing. These demonstrate that radio observations could significantly improve the systematics for next generation optical experiments, if robust methods of analysis are developed.

Motivated by the promising results discussed in the previous chapter I then turn to analysing the radio lensing signal in chapter 5. I present an analytical model of a radio observation illustrating how a lensing signal distorts the data in the Fourier domain. I go on to develop a method for extracting this signal called "FILM" that estimates lensing shear directly from the radio visibilities. Using simulations based on the semi-analytical galaxy models of Wilman et al. (2008) and the expected specification of an SKA1-mid survey, I test the FILM method on different cases with various levels of realism. I find that given a high quality data set this method is able to reproduce the initial shear signal with noise properties related to the intrinsic shape dispersion as expected. The method is also found to be robust against several other sources of noise that would be present in real data. Further work is needed in the normalization of this measure and there are many other physical parameters that could be included. However, this work demonstrates a proof of concept for direct radio weak lensing estimators.

Finally, in chapter 6 I discuss some other projects I have conducted that are related to the themes of this thesis and would benefit form further work in the future. I discuss the imaging of wide-field low frequency observations from LOFAR, early analysis of ellipticity measurements in the Boötes field and the potential of machine learning techniques for radio-optical cross matching. I finish with a summary and conclusions of the thesis.

1.5.1 Current constrains on ΛCDM parameters.

Table 1.1 contains the latest constrains on cosmological parameters from the various probes described above. They are all calculated based on a flat wCDM model.

Probe	Reference	Ω_m	w	H_0 ($\mathrm{km\,s^{-1}\,Mpc^{-1}}$)	σ_8
CMB	Planck (2014)	0.314±0.02	$-1.13^{+0.13}_{-0.1}$	67.4±1.4	0.834±0.027
SNe 1a	Betoule et al. (2014); Riess et al. (2016)	0.295±0.034	-	73.24±1.74	-
Weak Lensing	DES Collaboration et al. (2017)	$0.264^{+0.032}_{-0.019}$	$-1^{+0.04}_{-0.05}$	$65.6^{+1.5}_{-2.6}$	$0.807^{+0.062}_{-0.041}$
TDSL	Bonvin et al. (2017)	-	$-1.79^{+0.94}_{-0.49}$	$79.1^{+4.4}_{-4.2}$	-
Planck + BAO	Alam et al. (2016)	0.302±0.012	-1.05±0.06	68.8±1.4	-
Planck + TDSL	Bonvin et al. (2017)	0.3±0.02	$-1.38^{+0.14}_{-0.16}$	$79^{+9.3}_{-8.7}$	-
Planck + SNe 1a	Betoule et al. (2014)	0.307±0.017	-1.018±0.057	68.07±1.63	-
Planck +BAO + SNe 1a	Betoule et al. (2014)	0.303±0.012	-1.027±0.055	68.5±1.27	-

Table 1.1 Latest constrains on cosmological parameters from the various probes described above. They are all calculated based on a flat wCDM model.

Chapter 2

Radio Astronomy

This chapter will provide the underlying principles governing radio astronomy and the contribution it can make to the wider astronomy and cosmology fields. The vast majority of current and next generation cosmological radio instruments are arrays; as such this chapter also includes a description of interferometry and aperture synthesis which forms the motivation for the technique developed in chapter 5.

2.1 The Radio Window

For most of the history of astronomy our knowledge of the Universe has entirely comprised of what can be observed using visible light. In recent years the range of frequencies from electromagnetic radiation available for study has grown to span almost the entire spectrum. Today the ability for astronomers to further develop their understanding of the Universe is limited only by how much can be collected and analysed. Most advances and further constraints on cosmology have come from large sets of multi-frequency data where different physical processes can be probed and related or removed. Indeed the next generation of precision cosmology is even moving beyond the EM spectrum altogether, probing gravity itself (Abbott et al., 2016).

Figure 2.1 shows the opacity of the Earth's atmosphere over a range of EM wavelengths. Although some non-optical frequencies, namely the near infra-red and ultraviolet, can be observed using ground based telescopes, most require us to resort to expensive space based instruments. However, in 1931 Karl Jansky identified radiation at a wavelength of 14.6m, a previously un-probed frequency; in showing that the origin of this signal was galactic (Jansky (1935)) he opened a new window on the Universe. This transparent gap from $\lambda = 10^{-4}$m to 10m along with technical advances in receivers during World War II allowed us to observe

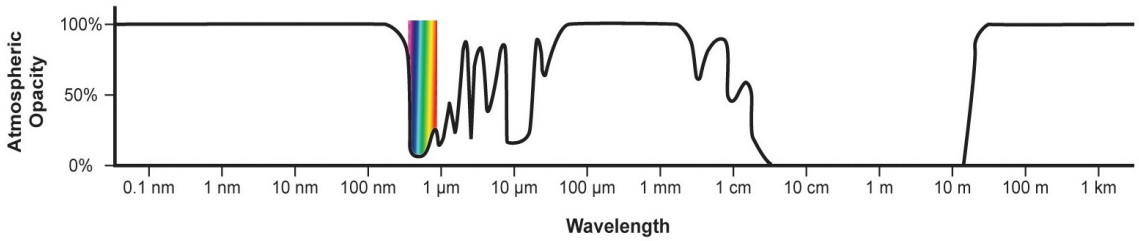


Fig. 2.1 Atmospheric opacity as a function of EM-radiation wavelength. (Image Credit: NASA, http://www.ipac.caltech.edu/outreach/Edu/Windows/irwindows.html)

entirely new physical mechanisms without the limitation of having to escape the Earth's gravity.

2.2 Single aperture telescopes

We begin by considering basic radio telescopes that only have one receiver; these produce a single instantaneous measurement of a patch sky. We will explore in later sections how these have become the individual elements of modern arrays increasing the efficiency of observation.

Antennas

We can consider a radio telescope as an antenna, a receiver (or transmitter) of a radio radiation field. The properties that govern the effectiveness of a telescope are its sensitivity and angular resolution. These qualities are dependent on the exact design and quality of the instrument but are fundamentally constrained by the effective collecting area that intercepts radiation. The simplest design for a radio antenna is a dipole that receives isotropically and would therefore observe the whole sky simultaneously. In this case the power received (P_ν) over some bandwidth ($\Delta\nu$) is simply proportional to the total flux of the radiation field (I_ν), and the response A_0.

$$P_\nu(\theta,\phi) = A_0 I_\nu(\theta,\phi)\Delta_\nu \tag{2.1}$$

where θ and ϕ define a spherical coordinate direction. Although a wide field of view can be useful, (see for example science cases cited by Lonsdale (2007) for the Murchison Widefield Array, MWA) the directional sensitivity can be increased by concentrating power

in a particular direction. If an antenna instead has a reception pattern given by the collecting area $A(\theta,\phi)$ (in m^2) with a maximum response of A_0, then the power received becomes,

$$P_\nu(\theta,\phi) = A_\nu(\theta,\phi)I_\nu(\theta,\phi)\Delta_\nu \tag{2.2}$$

The response pattern induced through the collecting area is known as the primary beam; the beam solid angle (Ω_A) can then be used as a measure of a telescope's field of view. This is given by Burke & Graham-Smith (2010) as,

$$\Omega_A = \int_{4\pi} \frac{A(\theta,\phi)}{A_0} d\theta d\phi \tag{2.3}$$

Due to the conservation of energy the peak reception and field of view are fundamentally linked by $A_0 \propto \Omega_A^{-1}$, hence as collecting area and sensitivity improves the beam solid angle must decrease.

In order to have a consistent measure for the brightness of an object, the received power is normalised by collecting area (m^2), unit of bandwidth (Hz), and steradian of sky. The total flux of a source is then given by integrating over the solid angle subtended, and hence is in units of $W\,m^{-2}\,Hz^{-1}$. Due to the typically low fluxes of radio sources the Jansky (Jy) is conventionally used as the unit of flux density, defined as,

$$1Jy = 10^{-26} W\,m^{-2}\,Hz^{-1} \tag{2.4}$$

Parabolic dishes

Typically the collecting area of a radio antenna is increased by providing a reflecting surface that concentrates the radiation at a single antenna. Designs for different systems of radio reflecting dish can be seen in Fig. 2.2. Each of these systems is based on a parabolic surface with various feed mechanisms; they all generate different reception and noise patterns, so the best choice can vary with science case. For example, the C-BASS telescopes (Pearson & C-BASS Collaboration, 2016) are designed to measure and remove the foreground emission from the CMB polarization data. Since these measurements are highly sensitive to angular distortions, C-BASS uses a prime focus reflector supported by radio transparent foam. This results in a lower overall sensitivity but a circularly symmetric beam response.

At long wavelengths secondary reflectors need to be very large, as such a prime focus system is generally used when a large frequency range is required. An exception to this is Arecibo which requires a secondary dish as the primary is spherical rather than parabolic.

Despite this, Cassegrain reflectors are generally more popular since they allow for a larger feed with easier access, faster frequency band switching, and in the case of offset systems, an unblocked aperture providing an almost circularly response and low side-lobes. Some other reflecting designs have been developed that use only portions of the parabolic surface, either to minimise noise like the Holmdel Horn Antenna, or maximize angular resolution e.g. the RATAN-600. However, from now on we will consider instruments based on dishes or dipoles.

The sensitivity of a dish can be increased by making the surface area larger; this is adequately described by the proxy of the diameter of dish D. As we have seen, with increasing diameter and surface area the field of view of an antenna shrinks. Since a single dish only produces one measurement, a larger diameter leads to smaller pixels. The best theoretical angular resolution for a dish of diameter D is limited by diffraction as $1.22 \times \lambda/D$, where λ is the central observing wavelength. The units used here are radians for resolution and meters for diameter and wavelength.

2.3 Interferometry

In the previous section we found that by increasing the collecting area of a telescope (D in the case of a dish) both the sensitivity and angular resolution could be increased. We also saw that this increase is related to the observing wavelength, which is much longer for radio than most of the rest of the spectrum. For example, the Hubble Space Telescope (HST) has a primary mirror with $D = 2.4m$; observing in visible light this gives HST a resolution of $\sim 0.05''$. To achieve the same detail with at metre wavelengths would require a dish thousands of kilometres wide.

Clearly building a parabolic reflector on these scales would not just be expensive but impractical to the point of impossibility to construct. As such, astronomers have combined the signal from multiple dishes or antennas into arrays to simulate a larger instrument. This technique is called interferometry and we now consider the processes involved in generating an image from multiple observations and the problems therein.

2.3.1 Spacial coherence of a radiation field

The fundamental principle of interferometry is that two measurements of a propagating field (radio waves in this case) taken at different locations will not necessarily contain the same

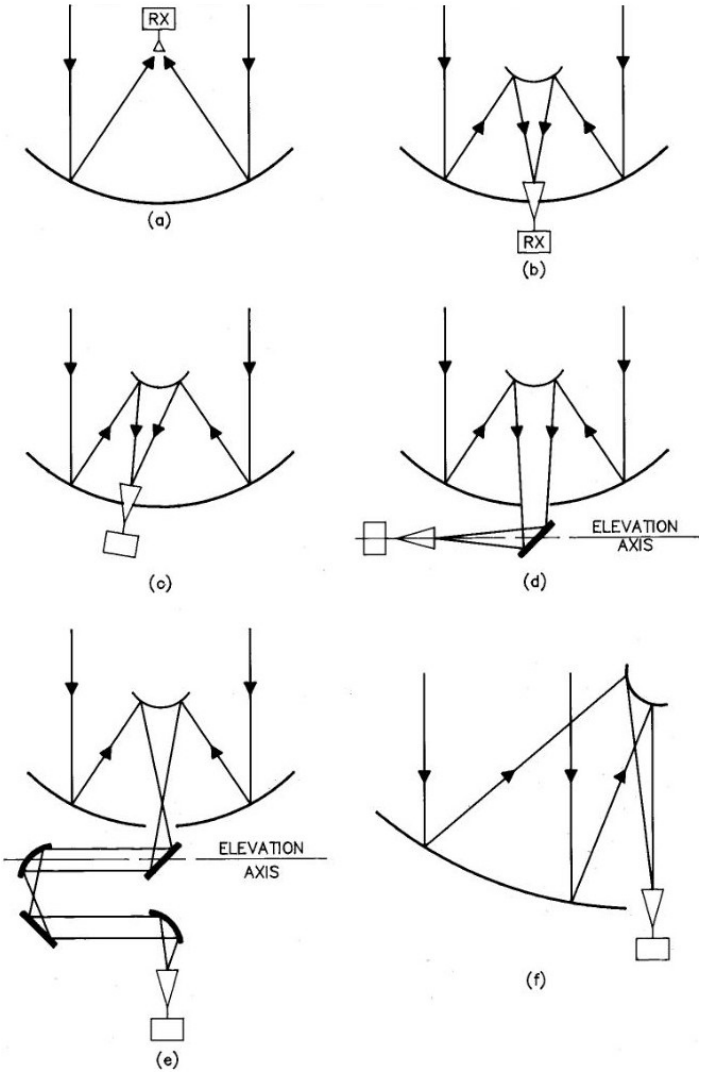


Fig. 2.2 Various systems of reflector for radio telescopes: (a) Prime focus, (b) Cassegrain, (c) Off-axis Cassegrain, (d) Naysmith, (e) Beam waveguide, (f) Offset Cassegrain. CREDIT: NRAO

information. The degree to which the field varies between these points can be described using the spatial coherence function V, which in turn is related to the source of the field itself. The coherence is simply measured by the correlation between a function and its complex conjugate i.e.

$$V(\mathbf{a},\mathbf{b}) = \langle f(\mathbf{a}), f^*(\mathbf{b})\rangle \tag{2.5}$$

For a physical field the parameters that the function f depends upon tend to be space and time. The mutual coherence function is a correlation taken with different values for both these parameters; if only time is varied then it is called the autocorrelation; spatial coherence is given by comparing two locations at the same time.

Consider a luminous astrophysical signal at some large distance producing the radiation field **E**. Since structure along the line of sight cannot be disentangled, the only observable quantity of **E** is its surface brightness distribution. This can be thought of as a projection of **E** onto some celestial sphere at distance **R** which shall be denoted by $\mathscr{E}(\mathbf{R})$. If there is no further contribution to the radiation field within this sphere, then its propagation is given by Huygens' principle (Taylor et al., 1999) as simply,

$$E_\nu(\mathbf{r}) = \int \frac{\mathscr{E}_\nu(\mathbf{R})}{|\mathbf{R}-\mathbf{r}|} e^{2\pi i|\mathbf{R}-\mathbf{r}|\nu/c} d\Omega \tag{2.6}$$

where $d\Omega$ is an element of surface area on the sky. This defines the electric field received by an antenna at position **r**. Taking the spatial coherence of this signal from two points $\mathbf{r}_1$ and $\mathbf{r}_2$ gives,

$$V_\nu(\mathbf{r}_1,\mathbf{r}_2) = \langle E(\mathbf{r}_1), E*(\mathbf{r}_2)\rangle \tag{2.7}$$

$$V_\nu(\mathbf{r}_1,\mathbf{r}_2) = \int \frac{|\mathbf{R}|^2\langle\mathscr{E}^2_\nu(\mathbf{R})\rangle}{|\mathbf{R}-\mathbf{r}_1||\mathbf{R}-\mathbf{r}_2|} e^{2\pi i|\mathbf{R}-\mathbf{r}_1|\nu/c} e^{-2\pi i|\mathbf{R}-\mathbf{r}_2|\nu/c} d\Omega \tag{2.8}$$

We can write this autocorrelation in terms of the observed intensity $I(\hat{s})$, where $\hat{s}$ is the unit vector $\mathbf{R}/|\mathbf{R}|$. The coherence function can also be simplified by making the assumption that,

$$|\mathbf{r}| << |\mathbf{R}| \Rightarrow \left|\frac{\mathbf{r}}{\mathbf{R}}\right| \simeq 0, \tag{2.9}$$

which is in general true for astrophysical observations. This results in,

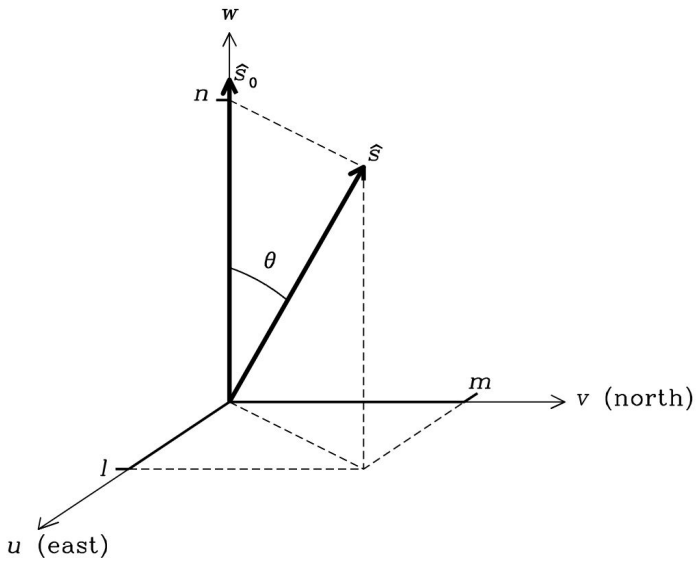


Fig. 2.3 The (u,v,w) coordinate system. CREDIT: Burke & Graham-Smith (2010)

$$V_\nu(\mathbf{r}_1,\mathbf{r}_2) \simeq \int I_\nu(\hat{s}) e^{-2\pi i \hat{s}|\mathbf{r}_1-\mathbf{r}_2|\nu/c} d\Omega \tag{2.10}$$

Notice that the measurement V is only a function of direction ($\hat{s}$) and separation rather than exact location of antennas ($|\mathbf{r}_1-\mathbf{r}_2|$). We call the vector between two antennas a baseline ($\mathbf{b}=|\mathbf{r}_1-\mathbf{r}_2|$); an interferometer will give one data point per baseline which is referred to as a visibility. The visibilities produced by an interferometer are not exactly the spatial coherence function required since they are also dependent on the properties of the particular instrument, as we will discuss later in this chapter. We can find the theoretical resolution limit of an interferometer by using its maximum baseline b_{max} as a proxy for the "diameter" of an equivalent dish i.e.

$$\theta_{max} = \frac{\lambda}{b_{max}} \tag{2.11}$$

2.3.2 The (u,v,w) coordinate system

A special coordinate system called (u,v,w) is used in interferometry in order to describe baseline positions and pointing direction relative to the geometry and motion of the Earth. This system uses three components based on a projection of Cartesian coordinates and can be seen in Fig 2.3. The w coordinate is defined in some reference direction ($\hat{s}_0$) for the observation, usually toward a particular source or the centre of the observing field. The (u,v) plane can then be constructed as being normal to w and pointing east and north respectively.

The baseline vector normalized by wavelength ($\mathbf{b}/\lambda$) can now be fully described as a combination of u,v, and w. The pointing position on the sky ($\hat{s}$) can be projected onto the plane of the sky with positions l and m in this space. Since $\hat{s}$ is a unit vector this dictates its w position to be $n = \sqrt{1-l^2-m^2}$, thus $l^2+m^2+n^2 = 1$. The direction $\hat{s}$ is also sometimes defined as a set of three angles (θ, γ, β) away from the (u,v,w) axes; these are related to (l,m,n) by cosine.

$$n = \cos\theta \tag{2.12}$$

$$l = \cos\gamma \tag{2.13}$$

$$m = \cos\beta \tag{2.14}$$

In this coordinate system the spatial coherence function in equation 2.10 becomes,

$$V_\nu(u,v,w) = \int I_\nu(l,m,n) e^{-2\pi i(ul+vm+wn)} dldmdn \tag{2.15}$$

$$\equiv \int \frac{I_\nu(l,m)}{\sqrt{1-l^2-m^2}} e^{-2\pi i[ul+vm+w(\sqrt{1-l^2-m^2}-1)]} dldm \tag{2.16}$$

The frequency term in the exponential has been absorbed into the coordinate system here since we defined baseline positions normalised by λ.

2.3.3 Fourier inversions of spatial coherence

In order to retrieve the sky brightness distribution from the spatial coherence function equation 2.16 must be inverted. To begin we examine two special cases that, whilst simplifying the equations, do not have an inversion procedure greatly different from real interferometric radio arrays. First we suppose that antennas all lie in a coplanar array such that baselines have no w component to their position, and second that the reception pattern is only sensitive to a small region around $\hat{s}_0$.

Since the separation of antennas in a radio array is often small compared to distances over which the earth curves, the w component of baselines will be neglected here. By setting $w \simeq 0$ in equation 2.16 we obtain,

$$V_\nu(u,v,w \simeq 0) = \int \frac{I_\nu(l,m)}{\sqrt{1-l^2-m^2}} e^{-2\pi(ul+vm)} dldm \tag{2.17}$$

This is clearly a Fourier transform and hence can be easily inverted to give the intensity of the sky in direction (l,m).

$$\left(1-l^2-m^2\right)^{-\frac{1}{2}} I_\nu(\hat{s}) = \int V_\nu(u,v) e^{2\pi(ul+vm)} du dv \tag{2.18}$$

The assumption that $w = 0$ does not hold in the case of Very Long Baseline Interferometry (VLBI), where Earth curvature is not negligible. Also, depending on the location and the direction of an observation even a coplanar array can shift into the w coordinate through rotation of the Earth. However, it has been shown that the general relationship between I and V is a still Fourier transform (Taylor et al., 1999). There is little difference between the two and three dimensional relationships other than the calculation being more time consuming. As such, in chapter 5 I only use a two dimensional transform. This would require generalising before an application on real data.

The second special case we consider is that of an observation restricted to a small region of sky. This case is more useful for modern observations where long integrations mean the array is rarely in a plane; we will also see how the result can be generalised for wide-fields. To define a small region of observation we restrict the pointing direction ($\hat{s}$) to be at most a distance of σ away from $\hat{s}_0$, where $\hat{s}_0$ and σ are perpendicular to first order ($\hat{s}_0 \cdot \sigma \simeq 0$). Thus the (u,v,w) coordinate system is such that $\hat{s}_0 = (0,0,1)$ and $\sigma = (l,m,0)$ and equation 2.16 becomes,

$$V_\nu(u,v,w) = e^{-2\pi i w} \int I_\nu(l,m) e^{-2\pi(ul+vm)} dl dm \tag{2.19}$$

The exponential term involving phase can be absorbed into the LHS of the equation. This gives an expression for V relative to $\hat{s}_0$ that is independent of w,

$$V_\nu(u,v) \equiv e^{-2\pi i w} V_\nu(u,v,w) = \int I_\nu(l,m) e^{-2\pi(ul+vm)} dl dm \tag{2.20}$$

Since the w component of $\hat{s}$ causes a phase change in the relative visibility, we call $\hat{s}_0$ the phase tracking centre. Again we obtain a Fourier transform for the form of V giving a direct inversion of,

$$I_\nu(\hat{s}) = \int V_\nu(u,v) e^{2\pi(ul+vm)} du dv \tag{2.21}$$

with a known phase dependence on θ. The additive phase difference for a baseline b pointing $\Delta\theta$ away from the phase centre is given by $2\pi\lambda b \cos\theta\Delta\theta$. For observations with a small fov the difference is calculated from the phase tracking centre and removed from the

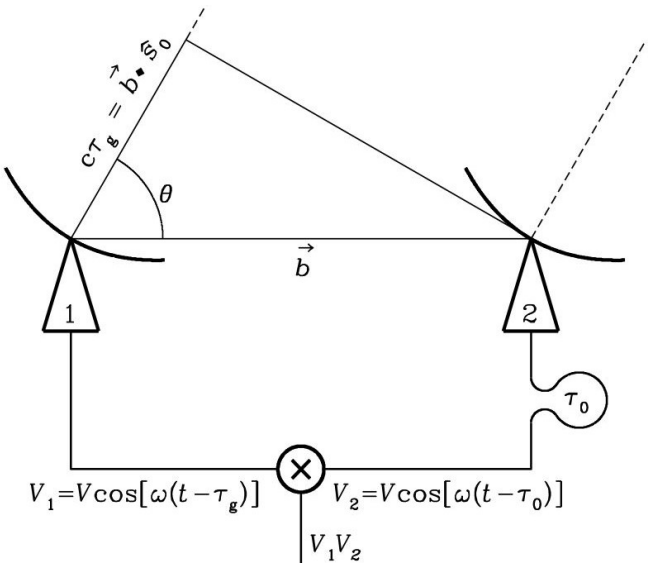


Fig. 2.4 A basic interferometer made from only two receiving elements separated by baseline b, pointing in direction $\hat{s}$. $\omega \equiv 2\pi\nu$ CREDIT: cv.nrao.edu/course/astr534/Interferometers1.html

observation. For wide-field observations however, this would cause excessive smearing at large $\Delta\theta$ so the imaging must be repeated in different directions with a new delay.

2.3.4 Two element interferometer

In this section I consider the simplest interferometer, one with only two elements, and hence one baseline, that produces a single visibility measurement. This is useful for understanding the process involved in combining measurements from array elements. At the end of this section I show that this simple array also results in equation 2.20. As the number of receivers (N_{ant}) increase the instrument becomes much more complicated since the baselines (N_b) grow as the square of the number of receivers,

$$N_b = \frac{1}{2} N_{ant}(N_{ant} - 1) \tag{2.22}$$

The set-up of an interferometer with two elements can be seen in Fig. 2.4; they are separated by a baseline **b** and pointing toward a compact source in direction ($\hat{s}$). A monochromatic radio signal arrives at antenna 1 after a delay of τ_g from antenna 2; this delay is caused by the extra distance travelled by the radiation and hence has the form,

$$\tau_g = \frac{\mathbf{b} \cdot \hat{s}}{c} \tag{2.23}$$

We can also induce a delay on the signal from antenna 2 (τ_0) to compensate for this, although for now we will consider the case where $\tau_0 = 0$. The voltage from antenna 1 therefore will simply have a phase shift induced by the delay, and the outputs can be described as,

$$v_2 = v\cos(2\pi\nu t) \tag{2.24}$$

$$v_1 = v\cos[2\pi\nu(t-\tau_g)] \tag{2.25}$$

The amplitudes from the two antennas are multiplied together by the correlator and then averaged in time such that the high frequency terms are cancelled out; this gives the cross-correlation output R as,

$$R(\tau)\langle v_1, v_2\rangle = \frac{1}{2}v^2\cos(2\pi\nu\tau_g) \tag{2.26}$$

The amplitude of the correlation is proportional to the flux of the source multiplied by the reception pattern as we saw in section 2.2. This pattern is given by combining the effective collecting area of the two receivers, $A = \sqrt{A_1 A_2}$. As the Earth rotates and the telescope tracks the source, the direction of observation ($\hat{s}$) changes. This causes the projected baseline length (and thus τ_g) to change and produces quasi-sinusoidal waves in the correlator output called fringes which can be described by the fringe phase Φ,

$$\Phi = \frac{2\pi b\cos\theta}{\lambda} \tag{2.27}$$

If the projected baseline $b\sin\theta$ is on the order of several λ or more, then Φ will be very sensitive to θ making the fringe phase an accurate measure of compact source position.

We can generalise the response of the correlator from a point-like source to a sky brightness $I(\hat{s})$ by treating the response from a region dR which is smaller than the resolution of the telescope as a point source and integrating over the field of view. This gives a response of,

$$R_c = \int A(\hat{s})I(\hat{s})\cos(2\pi\mathbf{b}\cdot\hat{s}/\lambda)d\Omega \tag{2.28}$$

This is called the cosine response R_c and is only sensitive to the symmetric part of a particular brightness distribution. To probe the anti-symmetric part we need the sine correlation R_s. This is found by introducing a time delay to one antenna that produces a phase change of $\pi/2$ radians.

$$R_s = \int I(\hat{s})I(\hat{s})\sin(2\pi\mathbf{b}\cdot\hat{s}/\lambda)d\Omega \tag{2.29}$$

Combined, the two correlations give the complex visibility V that is related to the spatial correlation function we saw earlier. These visibilities are defined as,

$$V(\hat{s}) = \sqrt{R_c^2 + R_s^2}e^{-i\tan^{-1}(R_s/R_c)} \tag{2.30}$$

Thus, substituting in the expressions for $R_{c,s}$ from , the form for a visibility of an arbitrary sky brightness then becomes,

$$V_\nu = \int A_\nu I_\nu(\hat{s})e^{-2\pi i\mathbf{b}\cdot\hat{s}/\lambda}d\Omega \tag{2.31}$$

In reality radio sources and antennas are not monochromatic and work within a fixed frequency range called the bandwidth. The observation is an integral over this range and this causes the amplitude of V to be damped by a sinc function of the time delay and bandwidth. We can compensate for this at the phase centre by setting the delay on antenna 1 to be $\tau_0 = \tau_g$. Since the error has only been corrected in the direction of $\hat{s}_0$ it will still be present in other directions. This defines a usable field of view in which the difference is small. This is the same relationship we encountered in equation 2.20,

$$V_\nu(u,v) = s\int A_\nu I_\nu(l,m)e^{-2\pi(ul+vm)}dldm \tag{2.32}$$

We can rewrite the last visibility equation using (u,v,w) coordinates to recover the expression in equation 2.20 and hence the same inversion properties. Hence, the two element interferometer returns a spatial coherence function normalised by the response pattern (A) which is dependent on the combination of the individual collecting area.

2.3.5 Interferometry at other wavelengths

Although used almost exclusively for radio astronomy, the first interferometer proposed by Michelson was designed to compare phases of optical light. There is no theoretical reason why information from instruments working at shorter wavelengths could not be combined in the same way. Practically however it is challenging. The much higher frequency of visible light for example would requires extremely high temporal resolution in order to correctly manipulate the phases. This is not only technically difficult to achieve but also would involve vast amount of data storage. However, by creating the required time delays using additional optical cable and combining the received beams directly, some interferometry has been

performed at infra-red wavelengths.

The usable field of view effect discussed above is amplified several orders of magnitude at high frequencies. These observations are therefore most useful for small scale studies e.g. the physical processes or astrometric positions of single stars. One such experiment is GRAVITY, an infra-red beam combining instrument on the (Very Large Telescope) VLT; for details see Eisenhauer et al. (2011). The goals of GRAVITY include testing general relativity by looking for deviations in the orbit of stars in the strong field around super massive black holes (SMBHs).

2.4 Aperture synthesis

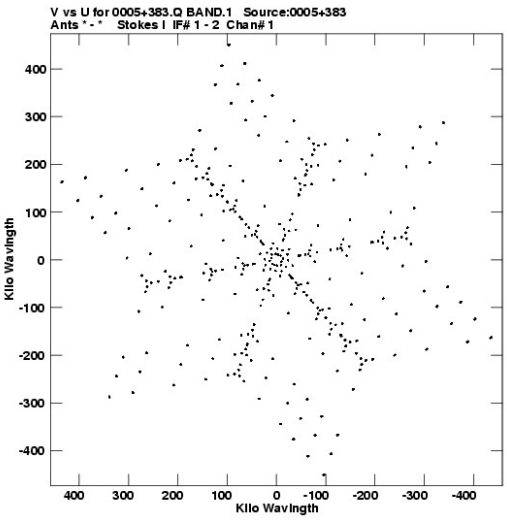
We have seen how the distribution of brightness on the sky can be sampled using the correlation between two antennas giving a complex visibility. The location of the visibility in (u,v,w) coordinate space is (l,m,$\sqrt{1-l^2-m^2}$) defined by the projected baseline. Additionally, we have seen that the relationship between visibilities and the true sky brightness is a Fourier transform (within certain limits). A single baseline however only defines one point in visibility space[1]. Using multiple baselines, or combining the same baseline with different projections allows us to produce a more complete data set. This technique is called aperture synthesis.

There are two ways in which aperture synthesis is performed, either by adding additional antennas, thus creating more interferometric pairs, or by allowing the rotation of the Earth to naturally change the projected baseline positions. Another method would be to manually move antennas to different separations. This has been done by combining various VLA configurations for example, but it would be impractical during an observation.

Increasing the number of receivers in an array to N creates N(N-1)/2 interferometric pairs. A large enough number of elements can produce an instantaneous uv coverage that is sufficient to reconstruct the sky at some level. A brief observation like this is referred to as a "snapshot" and the corresponding uv sample defines the characteristic response of the array. The classic "Y" layout of the VLA along with its snapshot uv coverage can be seen in Fig. 2.5, the six points of the star shape given by the longest baselines whilst the many short spacings fill the centre.

[1] Two actually, since the visibilities will have Hermitian symmetry, but only one unique point

(a) CREDIT: Google Earth



(b) CREDIT: National Radio Astronomy Observatory, *vla.nrao.edu*

Fig. 2.5 Left: the VLA seen from above showing the classic "Y" shaped array design. Right: Typical UV coverage of the VLA during a snapshot observation.

If the instantaneous sampling of the array is not sufficient then the rotation of the Earth can be used to fill more of the uv plane. Many short integrations are taken over a long period of time, calibrated for time variable effects, and then added to the same coordinate space as though they were measured coincidently. New arrays are now often designed such that they fill the UV plane as efficiently as possible for minimum rotation, shortening the required observation time.

2.4.1 Effect of discrete uv sampling

Despite the techniques used to measure as much of the uv plane as possible; the desire to have long baselines that provide high angular resolution makes a continuously sampled uv space practically impossible. As such the spatial coherence function V is sampled only at discrete locations, which can be described by the sampling function $S(u,v)$. This changes our visibility equation 2.20 to be,

$$S(u,v)V_\nu(u,v) = \int I^D_\nu(l,m)e^{-2\pi(ul+vm)}dldm \tag{2.33}$$

The modified intensity I^D_ν is called the dirty image and is related to the true sky intensity and the synthesised "dirty" beam or PSF B^D,

$$B^D(l,m) = \int B(l,m)S(u,v)e^{2\pi i(ul+vm)}dudv \tag{2.34}$$

through convolution,

$$I_\nu^D = B^D * A_\nu I_\nu \tag{2.35}$$

To reconstruct the true sky brightness from the sampled visibilities we need to deconvolve the PSF from the dirty image. Since the PSF of a radio interferometer is dependent only on the sampling function, which can be measured from baseline positions, it is theoretically known and deterministic.

2.4.2 Producing an image

There are many steps involved in creating an image from a set of array visibilities. These steps are complex, time consuming and highly non-linear. The problems in creating reliable images of high fidelity form part of the motivation for the work in chapter 5. I also look in more detail at creating and using images from LOFAR survey data in section 6. As such I give I a brief overview of some of the main steps below. However, since the majority of the work in this thesis is concerned with visibilities rather then images, I refer the reader to Taylor et al. (1999) for further details.

Gridding and weighting

When data is first taken with an interferometric array, it is not in a uniform pixel-like grid pattern. Rather, it is defined at non-uniform spacings, as seen in figure 2.5(b). In order to create an image, multiple Fourier transforms and convolutions need to be performed. In addition, for arrays with many elements, there will be redundantly over sampled regions of the *uv* plane, this creates large data volumes. As such this raw data is mapped onto an evenly spaced array to reduce the file size and increase computational efficiency. Unless the data is first convolved with a gridding kernel, this process in Fourier space can produce unwanted effects in the image. There are several gridding kernels that can be chosen with different levels of complexity. A typical kernel is a sinc function with the grid width defining the distance between first null points.

Once the data is successfully convolved it can be added to the grid. For grid pixels that contain multiple baselines a sum is taken and there are a few different ways that this can be

weighted. The most extreme methods are natural and uniform weighting. Natural weighting simply gives a constant weight to all visibilities, effectively down-weighting the signal from noisy measurements. This provides maximum point source sensitivity, but a poor beam shape and large side-lobes. Uniform weighting weights the data by the inverse number density of the sampling function. Basically, this means grid points are the average of the visibilities within the grid width. Uniform weighting can result in a much lower side-lobe contribution, but at the cost of higher noise. The weighting that is chosen is often somewhere between these two examples. Robust weighting, developed by Briggs (1995), provides an optimal trade-off for most cases.

So, just to get the data into a manageable form, two complex operations have had to be applied. Both of these use parameters that are decided by the observer and there is no consensus on the best option for either.

Calibration

Now that the data is in a vectorised format we can begin to calibration to reduce effects such as the distortion from the ionosphere. Radio Frequency Interference (RFI) from ground sources is also removed at this stage, using programmes such as AOflagger (Offringa et al., 2012) to identify unlikely spikes in amplitude. Calibration in radio works in much the same way as Adaptive Optics (AO) for optical instruments, although here radio does have an advantage. Alongside the main observation a well known compact object is also observed at intervals to measure changes in amplitude and phase, corresponding to brightness and position. In AO the calibration is performed during the observation and any time sensitive variation is lost. Conversely, for radio observations the calibrator data is stored for the corrections to be calculated and applied later. This means the calibration can be performed in a much more precise way, for example, applying solutions that vary with direction or removing particularly poor parts of the data. In fact for telescopes such as LOFAR, which is pointed through software rather than mechanically, calibrator sources can be observed at the same time as the main field for the whole observation. On the other hand, this does create another complex operation for the observer.

In chapter 6 I discuss the challenges of calibration in more detail, especially for wide field instruments. I also look at some novel solutions such as facet calibration (Williams et al., 2016). It should be noted that all of these techniques require multiple Fourier transforms and convolutions to calculate the corrections.

Deconvolution and imaging

Finally, the data is gridded and corrected and we are able to make an image. Because the visibilities were sparsely sampled, the relationship between the dirty and true image cannot be solved using a linear deconvolution and so a non-linear approach is needed. If we Fourier invert the gridded visibilities now we will obtain the dirty image, which is the true image convolved with the PSF or dirty beam as in equation 2.35. To reconstruct a likely sky brightness from the dirty image the degrees of freedom need to be reduced by incorporating a priori information. There are many different methods for performing this operation. One of these is called *CLEAN*, which has been the basis of the deconvolution technique used in almost all observations for the last 30 years. *CLEAN* is an iterative process which makes the assumption that the image can be represented as a sum of point sources in an otherwise empty FOV. The original *CLEAN* algorithm, developed by Högbom (1974), is shown below:

1. In the dirty image identify the position and strength of the peak in terms of absolute intensity;
2. Save this peak, multiplied by a user defined damping factor $\gamma \leq 1$, as point source, to a model. The point sources in this model are called the *CLEAN* components;
3. Subtract this *CLEAN* component, multiplied by a model of dirty beam, from the dirty image at the position of the peak;
4. Loop back to step (1) until the dirty image contains no peak above some user-specified level. The remainder of the dirty image is then termed the residuals;
5. Convolve the point source model with an idealised "*CLEAN*" beam. This is often an elliptical Gaussian fitted to match the primary response of the antenna elements;
6. The final *CLEAN* image is then the sum of the model in step (5) and the residuals from step (4).

Clearly this method is highly non-linear and user specific. Despite being widely used in some form for over three decades, the theory behind *CLEAN* and its possible distorting effects is still poorly understood. Reviews of *CLEAN* by Schwarz (1978); Wakker & Schwarz (1988) have shown the outcomes in various situations and identify that for certain parameters it can produce a unique solution. There have been many additions and improvements to the basic *CLEAN* algorithm since it was first created. However, for some science cases, even subtle unknown distortions would be enough to overwhelm the signal.

More recently there have been methods developed which handle deconvolution more robustly such as the Maximum Entropy Method (MEM), Compressed Sensing and UVMULTIFIT Cornwell et al. (1999); Garsden et al. (2015); Martí-Vidal et al. (2014). However, their use is not as widespread as *CLEAN* due to their relative youth.

Finally, it is important to note that for all these deconvolution methods the image produced is not a unique representation of the data. A radio image can only ever be a model of the visibilities, no matter how carefully it is performed. So, if one creates a model of a certain feature from a radio image, say a shear map, this is actually a model of a model of the data. Thus, radio observations need to be analysed carefully.

2.5 Radio galaxies

In this final section we consider the processes which lead to radio emission from galaxies; further details can be found in Condon (1992); Rohlfs & Wilson (2000). We also examine the population of extra-galactic radio sources in preparation for the work in chapter 3. Mainly we discuss the results of the s-cubed simulation by Wilman et al. (2008).

There are broadly two types of radio emission in the Universe, broadband continuum emission and line emission from transitions in atomic states. While the sharp and strong features of emission lines are useful for studying the dynamics of stellar environments, or galactic gas clouds, the work in this thesis is based on the radio continuum and we consider continuum-related processes only from now on.

The majority of emission at radio wavelengths is from synchrotron radiation caused by the acceleration of electrons to relativistic speeds in the presence of a magnetic field. Another important mechanism that causes radio waves is free-free or "bremsstrahlung" emission from regions of HII gas. At high frequencies thermal (IR) radiation from dust and stars begins to dominate the signal and the atmosphere also becomes less transparent as seen in Fig. 2.1.

The spectrum of the radio galaxy M82 is shown in Fig. 2.6, along with the contribution from these emission processes. At very long wavelengths most radiation is obscured by the Hydrogen gas in the galaxy. At frequencies above 100MHz the synchrotron component can be seen to follow a power law decreasing with frequency. Free-free emission begins to dominate above $\sim$ 30GHz as it has a shallower spectral index. It continues to be the main

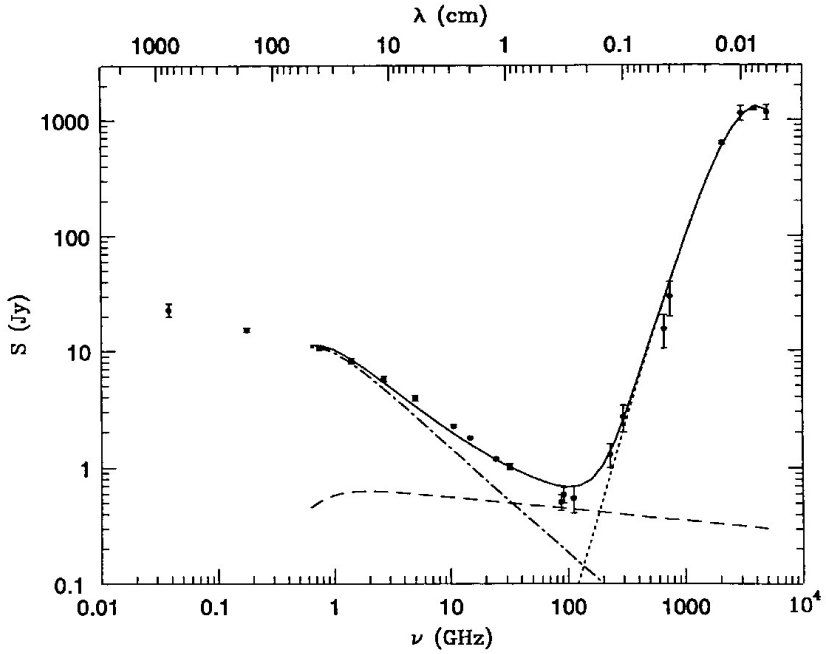


Fig. 2.6 The radio/FIR spectrum of M82 as seen in Condon (1992). The solid line is the sum of the other three components, synchrotron radiation (dot-dash line), free-free emission (dashed line), and a thermal component from dust (dotted line).

source radiation up until the IR regime.

2.5.1 Synchrotron radiation

This type of emission process was discovered from the construction of the first electron synchrotrons, hence the name. It was found that charged particles moving at relativistic speeds through a magnetic potential would lose energy as radiation. Since magnetic fields and charged particles (mostly electrons) are abundant in the Universe we expect to see this radiation from galaxies. The power and lifetime of synchrotron radiation in a particular galaxy is dependent on the magnetic field strength and relative electron velocities which can be predicted from lab based results.

An electron with energy E GeV, moving at angle θ to a magnetic field of strength B μG will emit mostly around a critical frequency given by,

$$\nu_c(\mathrm{GHz}) \simeq 1.6 \times 10^{-2} B \sin\theta E^2 \tag{2.36}$$

Typical magnetic field strength in spiral galaxies is $\sim 5\mu G$; this gives a synchrotron emission frequency of $\nu_c \sim 0.1E^2$GHz. So for these electrons to produce a signal in the accessible frequency range (0.1-10GHz) they must have energies of around 1-10GeV. The rate of energy emission in $\mathrm{GeV\,s^{-1}}$ is given by Larmor's relativistic formula as,

$$\dot{E}(\mathrm{GeV\,s^{-1}}) = -3.8 \times 10^{-18}(BE\sin\theta)^2 \tag{2.37}$$

Clearly even for strong fields, and electron energies of hundreds of GeV this is a very slow loss rate and synchrotron field lifetimes will be very long. By assuming an isotropic velocity dispersion for θ Condon (1992) calculates this as,

$$\tau_s(yr) = \frac{E}{\dot{E}} \simeq 1.06 \times 10^9 \sqrt{(B\sin\theta)^{-3}\nu_c^{-1}} \tag{2.38}$$

In the case of $\nu_c = 1.5$GHz, and $B = 5\mu$G this gives an estimate for τ_s of $\sim 10^8$ years. The power spectrum of synchrotron radiation as seen in Fig. 2.6 follows a power law where the gradient is given by a spectral index p. This allows us to study the underlying distribution of electron power in the Universe. The power from a single electron at frequency ν is given by Rohlfs & Wilson (2000) as,

$$P(\nu) = \frac{\sqrt{3}e^3 B\sin\theta}{m_e c^2}\frac{\nu}{\nu_c}\int_{\nu/\nu_c}^{\infty} K_{5/3}(\nu')d\nu' \tag{2.39}$$

The integral of the Bessel function K is simply a uniform square over the range $[\nu/\nu_c, \infty)$ and thus, assuming isotropy for θ, is directly dependent on the energy distribution. We observe the quantity $N \cdot P(\nu)$ from both radiation and cosmic rays to be a power law. Since the received number of particles (N) should be independent of energy this implies a power law energy distribution. We describe a galaxy's synchrotron radiation distribution using its spectral index p as,

$$I(\nu) \simeq \nu^{\frac{1}{2}}(1-p) \tag{2.40}$$

The main physical processes that cause synchrotron radiation are the large amounts of energy released by supernovae from stars with mass $> 8M_\odot$, and the accretion of matter onto a super-massive black hole (AGN). The second of these causes radio jets and lobes often at large distance from the galaxy itself; these are referred to as Fanaroff-Riley (FR) type or "classical" radio galaxies.

2.5.2 Free-free emission

As electrons pass through an area of ionized gas, electrostatic interactions between the two cause an acceleration of the electron which produces a photon at a radio wavelength. The low energy of a radio wave, compared to the large kinetic energy of an electron, means that the deviation to the electron's path is small. This path deviation can be approximated by a small perpendicular acceleration from which Rohlfs & Wilson (2000) define the total emitted energy as,

$$W = \frac{\pi Z^2 e^6}{4c^3 m_e^2 p^3 v} \tag{2.41}$$

where Z is the charge of the ion particle in elementary charge units, p is the minimum distance between the two particles called the impact parameter, and v is the velocity of the electron. During the "collision" time ($\tau = p/v$) a photon is produced in a single pulse with an almost uniform power spectrum over a range of ν. There is a cut-off frequency ν_g above which there is very little emission, as seen in Fig. 2.6. Rohlfs & Wilson (2000) calculate this cut-off to be,

$$\nu_g(\mathrm{Hz}) = \frac{1}{2\pi\tau} = \frac{3\pi v}{32p} \tag{2.42}$$

Free-free radiation from a single electron-ion interaction will depend on the dynamics, ratios and temperature of the two components. Emission from a whole region of a galaxy can be described as an average for that region, and is given by the free-free emission coefficient,

$$\varepsilon(\nu) = \frac{4Z^2 e^6 N}{3c^3 m_e^2}\sqrt{\frac{2m_e}{\pi kT}}\ln\bar{p} \tag{2.43}$$

where N is found by multiplying the number densities of electrons and ions, T is the electron temperature and $\bar{p}$ is the mean of the minimum and maximum impact parameter.

To generate free-free emission electrons must pass within a small distance of a charged particle e.g. dense HII regions. Hydrogen in a galaxy is efficiently ionized in the presence of massive young stars; this makes free-free emission a good proxy for the rate of star-formation.

2.5.3 Radio galaxy populations

The population of extragalactic radio sources can broadly be split into two groups, defined by their main emission mechanism; active galactic nuclei (AGN) and lower power star-forming galaxies. The number counts for both these populations and the relationship to flux den-

sity has been well documented for bright sources ($\geq 10^3 \mu Jy$) in the GHz regime. There is good agreement between various radio surveys (Bondi et al., 2003; Ciliegi et al., 1999; Prandoni et al., 2001; White et al., 1997) that the differential source counts normalised by $S^{2.5}(Jy^{1.5}sr^{-1})$ fit a power law with an exponent of ≈ 2.7 at $S = 100 - 200\mu Jy$ rising with flux before reaching a peak at $1Jy$. This radio bright GHz sample is well studied, because not only are they easier to detect in the radio, there is also a rapid drop off in the brightness of the optical counterpart for μJy and low frequency populations (Waddington et al., 2001). This decreases the completeness of any follow-up spectroscopic survey needed to separate the source types.

The population of radio sources with $10^3 \mu Jy$ flux and below has been simulated using extrapolated radio luminosity functions by Wilman et al. (2008); this can be seen in Fig 2.7. This work suggests that source counts should continue to decrease with flux, driven by a fall in AGN sources but slowed by an increase in star-forming galaxies which then dominate the population. The dominance of star-forming galaxies is predicted at around $S_{1.4} = 100\mu Jy$, evidence for which can be seen in Simpson et al. (2006) although the flattening described there is more severe than expected. There is also a large variation between other surveys at low fluxes such as Seymour et al. (2004) and Hopkins et al. (2003) which deviate by a factor of 2 at $200\mu Jy$. For the most part the GHz population is well understood with good agreement between the models and data; low frequency low flux sources however are less well constrained.

The increase in star-forming galaxies described above is also expected to occur at longer (metre) wavelengths, however the change in dominant source type is projected at an order of magnitude higher flux ($\sim 10^3 \mu$Jy) and number density. Hence, it should be much easier to measure the source count curve. However, due to the reciprocal relation of frequency and resolution many small sources are below the limit of low frequency surveys. A Giant Metre-wave Radio Telescope (GMRT) survey by Sirothia et al. (2009) shows a clear flattening of the source counts below a thousand μJy at 325MHz, as expected from the Wilman et al. (2008) model. More recently a Low Frequency Array (LOFAR) survey at 150MHz by Williams et al. (2016) showed this flattening trend at a few thousand μJy.

The conformity seen in these surveys to the model suggests that star-forming galaxies do indeed play a significant role in the μJy source population. This is supported by Fomalont et al. (2002) who suggest that the steep spectral index of radio galaxies below 100μJy indicate a very low proportion of AGN. However, using a combination of multi-wavelength, morphological and spectral data Padovani et al. (2007) show that spectroscopic follow up

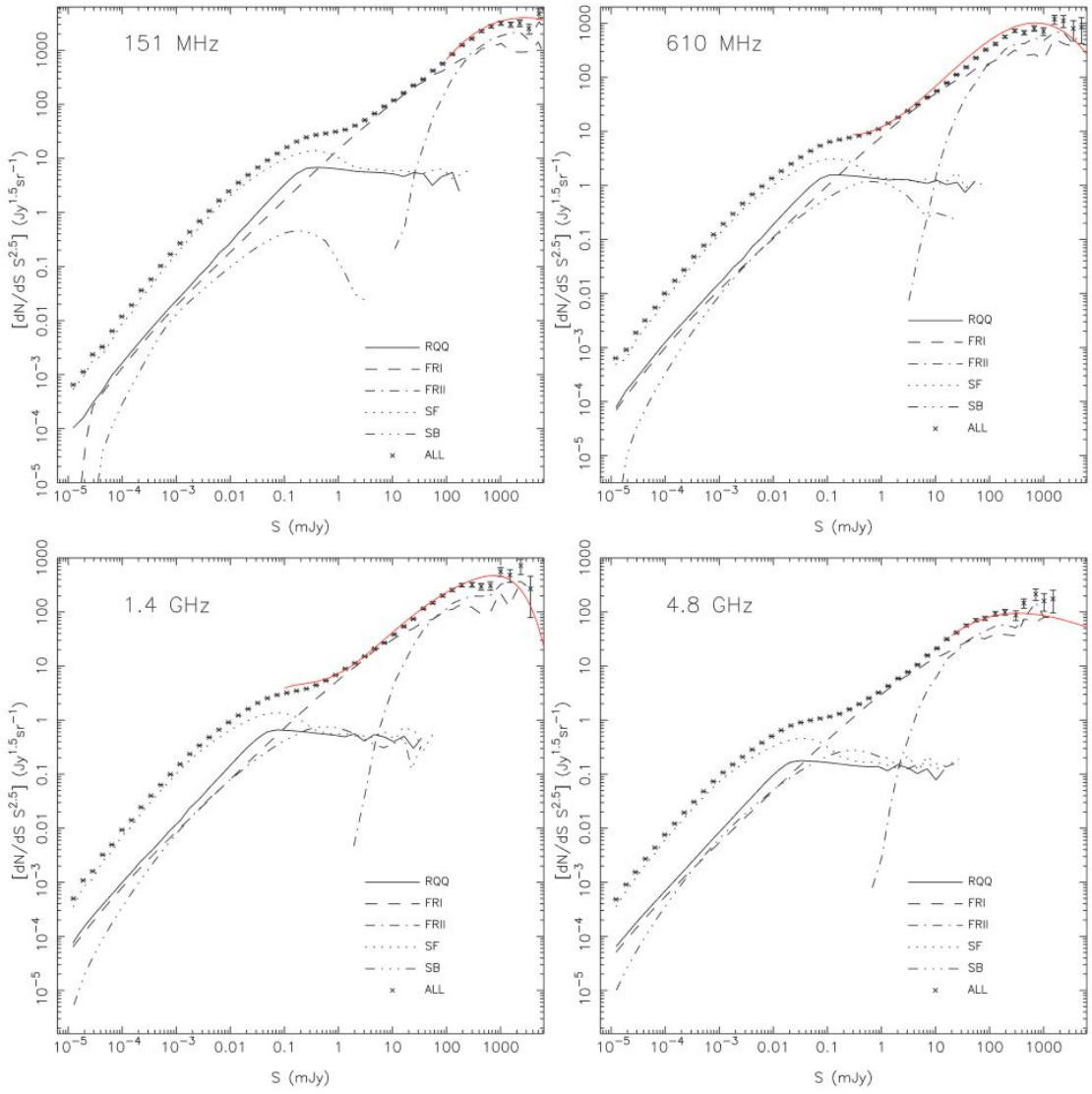


Fig. 2.7 Radio galaxy source counts at four different frequencies as modelled by Wilman et al. (2008).

can be biased toward a higher proportion of star-forming galaxies if the spectrograph is limited to magnitudes brighter than B=22 as early studies were (Benn et al., 1993). Indeed, Padovani et al. (2007) find that a minority (20-45 percent) of their sources below $10^3 \mu$Jy were star-forming.

Clearly, there is still significant work required to fully understand the population of μJy radio sources, especially at metre wavelengths. Such information is important for a range of astrophysical and cosmological studies. For example, the evolution of radio-quiet AGN, and their host galaxy, can constrain models of AGN feedback (Best et al., 2005), whilst the insensitivity of radio waves to dust obscuration assists studies of star-formation in distant galaxies (De Young, 1989). For cosmology, quantifying the radio source population will be critical for interpreting studies of the angular clustering of radio galaxies on large scales (Blake et al., 2004), and gravitational lensing measurements with the next generation of radio surveys (Harrison et al., 2016).

An next-generation survey that will provide some of this information is WEAVE-LOFAR, (Smith et al., 2016). This will be a spectroscopic follow-up of $\sim 10^6$ radio selected objects from LOFAR. Using ancillary time on SDSS I led a pilot programme to investigate the efficiency and accuracy of the standard BOSS pipeline for a radio population. This work is presented in the following chapter.

Chapter 3

Spectra of Faint Radio Sources in the ELAIS-N1 Field

The work described in this chapter has been conducted in collaboration with Robert C. Nichol, David J. Bacon, J. Sabater, N. Maddox, A. Kapinska, K. McAlpine, P.N. Best, A. R. Taylor and is based on Tarr et al. (2017a, in prep).

The radio source population has an impact on the radio cosmic shear correlation (Harrison et al., 2016) and studies of the angular galaxy clustering on large scales (Blake et al., 2004). Quantifying this population and hence its impact on such measures is essential for precision radio cosmology. The WEAVE-LOFAR experiment (Smith et al., 2016) aims to improve our understanding of radio galaxies using a spectroscopic survey of $\sim 10^6$ radio selected sources from LOFAR surveys. I led a team that was responsible for a pilot programme of this survey using current LOFAR and new BOSS data.

In this chapter, I report on the spectroscopic observations, presented in Tarr et al. (2017a), of a population of faint radio sources in the ELAIS-N1 field. These targets were selected from three, complementary radio samples spanning 150MHz-5GHz. In section 3.1, I describe these radio data including some of the first wide-field radio survey observations by the LOFAR telescope (Montes 2017 in prep), and deeper data (over a smaller area) from the upgraded Karl G. Jansky Very Large Array (JVLA) and GMRT. In section 3.2, I discuss the targeting and subsequent wide-field spectroscopic observations of thousands of radio sources.

In order to fully characterise the faint MHz radio population, a high resolution low frequency instrument which can integrate down to a 100μJy noise limit, coupled with an

efficient high throughput spectrograph is required. This is the design of the WEAVE-LOFAR survey (Smith et al., 2016) which aims to observe over 10^6 spectra of low-frequency selected radio sources with a high success rate. This chapter is a trial which aims to inform the targeting, cross matching and expected efficiency of the WEAVE-LOFAR survey by way of an ancillary programme in the Baryon Oscillation Spectroscopic Survey (BOSS; Dawson et al. 2013), part of the Sloan Digital Sky Survey (SDSS-III; Eisenstein et al. 2011).

In section 3.3, I provide details of the joint catalogue of sources which is publicly available as part of Tarr et al. (2017a) on which this chapter is based. Section 3.4 presents results of the spectroscopic follow-up and its success rate. A visual analysis of the spectra is provided in section 3.5 allowing us to consider the accuracy of the BOSS redshift pipeline. In section 3.6 I discuss the population of radio sources observed here combining the best data from visual inspection and the BOSS pipeline. Conclusions are given in section 3.7.

This was a collaborative effort and as such I was not directly involved in all aspects of the analysis. I did not carry out any of the observations used in this chapter radio, optical or spectral. I was also not responsible for the radio data reduction and imaging. However, I was part of the LOFAR "Blank Fields" working group which discussed the best analysis methods for each field, including ELAIS-N1. The matching of LOFAR sources with SDSS counterparts for targeting was carried out jointly by A. Kapinska and me. These radio-optical pairs were not preserved in spectroscopic results, and some sources could be found in multiple radio surveys. I was responsible for re-matching the spectra back to all radio objects for the final catalogue. I performed all analysis presented past section 3.3, with the exception of the visual redshift identifications, which as stated are the work of N. Maddox.

3.1 Radio Observations

In this section I report the radio observations used for this work. In particular I describe the calibration and source detection for the LOFAR observations which form the motivation for this study.

3.1.1 The ELAIS-N1 Field

The European Large Area ISO Survey - North 1 (ELAIS-N1) field is located at RA 16h14m00s and Dec +54:30:00 in an area of the galactic halo that contains no bright radio sources (>2Jy).

This provides low foreground obscuration of distant galaxies as needed to study galaxy populations over a range of redshifts. The field also possesses significant multi-frequency data as summarised in Table 3.1 and has been the focus of numerous studies of galaxy evolution and formation e.g. Arnalte-Mur et al. (2014); Franceschini et al. (2005); Garn et al. (2009); Kim et al. (2014); Manners et al. (2003); Pascual et al. (2001). The ELAIS-N1 field is also one of the most promising areas for the first detection of a signal from the Epoch of Reionization (EOR) from LOFAR observations (Vrbanec et al., 2015). Therefore, the ELAIS-N1 field was an obvious location for one of the first LOFAR galaxy surveys (see below) and is ideal for characterising the galaxy population in preparation for the next generation of wide-field radio surveys.

Instrument/ facility	Survey	Wavelength coverage	Area deg^2	References
Westerbork	WENSS	92 and 49 cm	570.0	Rengelink et al. (1997)
Herschel	HerMEs	100-500 μm	41.0	Oliver et al. (2012)
SIRTF	SWIRE	3.6-160 μm	8.7	Lonsdale et al. (2003)
Spitzer	SERVS	3.6 and 4.5 μm	2.0	Mauduit et al. (2012)
UKIRT	UKIDSS-DXS	1.17-1.33 and 2.03-2.38 μm	8.75	Lawrence et al. (2007)
Hyper Suprime-Cam	HSC-Deep	0.4-1.0 μm	7.2	Miyazaki et al. (2012)
SDSS	DR9 + BOSS	3000-10600 Å	14,555	Ahn et al. (2012)
GALEX	All sky survey	0.135-0.28 μm	20.0	Bianchi & GALEX Team (1999)
Chandra	EDXS	0.15-2.48 nm	0.2	Manners et al. (2003)

Table 3.1 Surveys of the ELAIS-N1 field showing a wealth of high quality data across multiple frequency bands.

Survey	Frequency	Bandwidth	Pointing centre α, δ (Deg)	Area (deg^2)	RMS noise (μJy beam^{-1})	Resolution	No. 5σ detections
LOFAR	150MHz	75MHz	243.5, 54.5	28	333	12"	800
FIRST	1365/1435MHz	21MHz	243.5, 54.5	28	145	5”	2617
GMRT Garn	610MHz	32MHz	242.75, 55.0	9	40/70	6”x5”	2500
GMRT Taylor	612MHz	32MHz	242.625, 54.5833	1.2	10.3	6.1”x5.1”	2800
JVLA	5GHz	2GHz	242.625, 54.5833	0.13	1.05	2.5”	483

Table 3.2 Proprieties of ELAIS-N1 radio observations used in this work.

3.1.2 LOFAR

The LOw Frequency ARray (LOFAR) is a next-generation, software-driven, international radio interferometer operating in the frequency range of 10-240MHz, where there has been little previous investigation. LOFAR consists of a core of 18 antenna stations in the Netherlands with long-baseline stations spread across Europe to provide the unique combination of a large instantaneous field-of-view (FOV), excellent angular resolution, and relatively high sensitivity. LOFAR also utilizes a novel phased array design in which signals are combined from many interferometric dipole antennas. Each of the dipoles in the LOFAR stations has a multi-directional response thus providing coverage of the whole sky. The telescope has no moving parts and therefore “points" to an object in the sky using digital delay between the stations, rather than moving a telescope dish as in a traditional radio observatory. Such software allows LOFAR to rapidly switch between targets and simultaneously observe multiple areas of the sky, which is particularly useful for monitoring calibration sources to correct for any time-dependent effects. LOFAR is also a key SKA-low pathfinder facility particularly for the EOR and cosmic magnetism working groups (Beck, 2007; Jarvis, 2007; Pritchard et al., 2015). The reader is referred to van Haarlem et al. (2013) for further details of the LOFAR instrument, and its major hardware and software components.

LOFAR observations began in December 2012 and the telescope is now collecting data on a number of Key Science Projects (KSPs) including a range of surveys of the northern sky (Rottgering, 2010). The LOFAR Tier-1 survey is the widest such survey covering the whole visible sky above LOFAR using both low-band (LBA; 15-65MHz) and high-band (HBA; 110-180 MHz) observations. The highest frequency band will reach an rms of 70μJy. There are also smaller area surveys with high quality multi-wavelength data: the Tier-2 "deep" survey which will reach 15μJy over 550 degrees^2 in its HBA band, and the Tier-3 "ultra-deep" survey, 5 fields observed using only the HBA covering 83 degrees^2 down to an rms of 7μJy at 150MHz.

ELAIS-N1 was selected to be part of the first cycle of LOFAR operations (cycle 0 February to November 2013) and was a major priority for many of the LOFAR KSPs. The observations were centred on $RA = 16^{\rm h}11^{\rm m}00^{\rm m}$, and $DEC = 55^{\circ}57'00''$ and received 120 hours of integration in the HBA data (115 to 190 MHz) reaching an rms depth of $S_{150\rm MHz} \sim 333\mu Jy$, over a field of view of $\sim 9\rm deg^2$, with an angular resolution of 12 arcsec.

The main ELAIS-N1 target field was observed simultaneously with 6 flanking fields including one centred on the compact radio source 87GB 160333.2+573543, for initial

calibration. The first dataset used is an eight hour observation, with the frequency coverage distributed evenly between the seven fields. After cleaning radio-frequency interference (RFI) from the data using the AOflagger (Offringa et al., 2012), the solutions from the calibrator field (containing 87GB) were transferred to the main ELAIS-N1 field. These solutions were based on a model of 87GB from data in several radio catalogues that characterize its flux, spectral index slope and curvature (Becker et al., 1991; Cohen et al., 2007; Condon et al., 1998; Douglas et al., 1996; Gregory & Condon, 1991; Hales et al., 1995; Rengelink et al., 1997). The LOFAR Black Board Selfcal (BBS) system (Pandey et al., 2009) was used to calculate and apply the calibration with AWimager (Tasse et al., 2013) used to produce an initial image. This was performed at seven different frequencies each with a bandwidth of $\sim$ 11MHz giving seven initial images. The distance between the main ELAIS-N1 field and the calibrator field is 2.6 degrees.

Sources extracted from these initial images were cross matched across frequency providing a high quality sky model for the main EN1 field including spectral index for some objects. This model was used to calibrate a second eight hour observation where more of the frequency coverage was used on the main field. This approach gives a better signal-noise ratio (SNR) than a standard one step calibration method.

It is important to note that due to the large field of view, relatively long baselines, and low frequency, there are direction dependent effects produced by the ionosphere (Intema et al., 2009) that cannot be corrected with the direction independent approach described before. The main effects on the catalogue are an increase on the background noise level from a theoretical level of $\sim 120\mu$Jy to $\sim 333\mu$Jy, and small shifts on the position of the sources which were estimated to be of the order of a couple of arc-seconds. There are also some inaccuracies in the flux level expected from the discrepancy between the real and the direction-independent solutions applied to each source. However, at the time that the catalogue was produced (March 2014), the newly developed, and more accurate, direction dependent calibration strategies were still being investigated (Tasse, 2014; van Weeren et al., 2016). Hence, the best calibration strategy available at the time was used.

A catalogue of sources was created using two detection tools, Python Blob Detection and Source Measurement (PyBDSM) which was developed for LOFAR images (Mohan & Rafferty, 2015), and BLOBCAT which was designed for the analysis of large survey data at radio frequencies in general (Hales et al., 2012). Both algorithms assign Gaussian islands of pixels to represent sources and link any islands which are likely to be part of the same

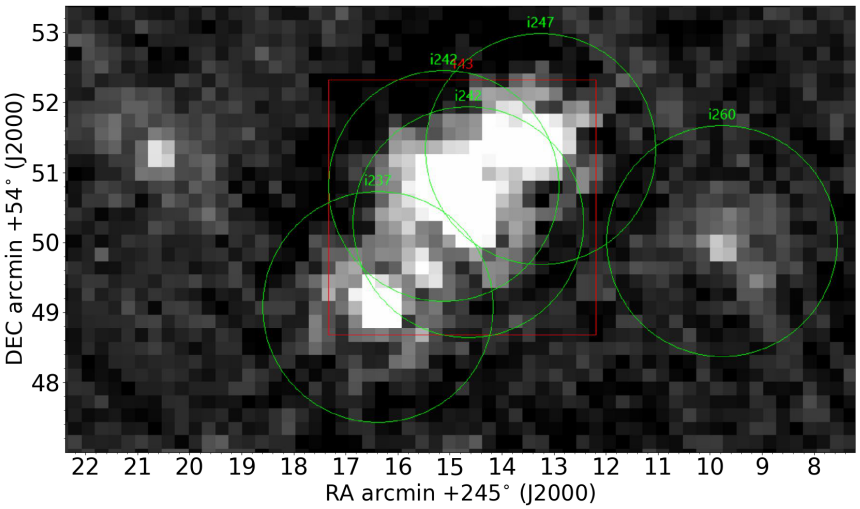


Fig. 3.1 LOFAR ELAIS-N1 cut-out centred at $\alpha = 245°11m30s$, $\delta = +54° : 50' : 30''$ showing a multiple component source and a mismatch between the PyBDSM and BLOBCAT detections. Green circles are PyBDSM islands, red square is from BLOBCAT.

extended structure, using noise maps to set detection limits for faint objects.

Using a 5σ threshold PyBDSM detected ~ 800 sources; lowering this to 3σ resulted in an additional ~ 400. Fainter sources which fell below the 3σ threshold were examined by eye, 387 of which were accepted resulting in a total of 1631 objects identified using PyBDSM. Conversely, BLOBCAT identified 942 unique source. This mismatch mainly occurs where BLOBCAT assigns only one island to multiple PyBDSM detections. This indicates that there are multiple source near each other, or an object has multiple components of emission. These mismatches in the source catalogues were used as one method for identifying multiple component objects. An example of this can be seen in Fig. 3.1.

3.1.3 VLA

In addition, we use sources from the FIRST catalogue and recent observations from the JVLA. First sources were chosen only in the area cover by the LOFAR EN1 field adding objects that may have been missed at low frequencies. The JVLA data is much deeper and concentrated in a small patch in the centre of ELAIS.

FIRST

Faint Images of the Radio Sky at Twenty Centimetres (FIRST; Becker et al. 1995) is a $\sim 10,000\,\mathrm{deg}^2$ survey of the north Galactic cap. It was conducted with the Very Large Array (VLA) in two narrow bands at 1365 and 1435 MHz from 1995 to 2010 and then a single frequency band centred on 1335 MHz in 2011. These maps have a typical rms noise of 150μJy, a resolution of 5", and sub-arcsecond positional accuracy.

The survey area of FIRST was chosen to coincide with the footprint of SDSS in order to encourage cross wavelength projects. Ivezić et al. (2002) used SDSS, sensitive at the time to a magnitude of 22.2 in r-band, to find optical counterparts to the FIRST sources. They were able to identify a match, defined by positional association within $1.5''$, for 30 percent of the FIRST catalogue.

JVLA

Due to upgraded receivers and a new correlator system the JVLA features 10 fold improvement in point source continuum sensitivity over the VLA. The JVLA observed 0.13 sq. degrees of the ELAIS-N1 field centred at $\alpha = 16^{\mathrm{h}}10^{\mathrm{m}}30^{\mathrm{s}}, \delta = +54°35'00''$ in 2012 and 2013. Observations were performed over 60 hours in 10 pointings using the "hybrid" B and C configurations at 5GHz. The final image was calibrated using radio sources 3C286 and J1624+5652 reaching an rms noise of 1.05μJy/beam, angular resolution of 2.5" and sub arc-second positional accuracy (Taylor et al., 2014).

3.1.4 Giant Metre-wave Radio Telescope

The Giant Metre-wave Radio Telescope (GMRT) operates in six frequency bands centred around 50, 153, 233, 325, 610 and 1420 MHz. It has a similar configuration, resolution and sensitivity to the JVLA with approximately three times the collecting area. Radio frequency interference (RFI) can be an issue for GMRT observations particularly in the higher frequency bands, but on average data quality is competitive with that of the JVLA.

In 2011 to 2013 Taylor et al. (2014) observed 1.2 sq. degrees of ELAIS-N1 centred on the same area as the JVLA field above. This observation was covered by a mosaic of 7 pointings each integrated for 30 hours at 612MHz with 32MHz bandwidth. The radio source 3C286 was again used for calibrating this data resulting in an rms noise of 10.3μJy/beam, with 6.1" x 5.1" angular resolution.

This field was also observed using the GMRT in 2003 for 25 hours (centred at $\alpha = 16^{\rm h}11^{\rm m}00^{\rm s}, \delta = 55^{\circ}00'00''$). A total of 9 sq. degrees were observed at 610MHz using a mosaic of 4 deep and 15 shallow pointings. The final image has a resolution of $6'' \times 5''$ and reaches an rms noise of 40μJy/beam in the deep areas, falling to 70μJy/beam elsewhere. For more details please see Garn et al. (2007, 2008).

3.2 Optical and spectroscopic observations

In this section I describe the optical observations used for creating a target list for subsequent spectroscopy. I discuss cross-matching of radio sources to known optical objects. In particular, I described the matching of multiple component radio objects which, in extreme cases, could be separated from their counterpart by several arc-minutes. This is a problem I investigate further in chapter 6. I give an overview of the spectrograph observation and data reduction performed by the BOSS team, describing the final target distribution and processing methods. Finally, I present details of the resultant catalogue and an analysis of the efficiency of the observation.

3.2.1 SDSS and BOSS

The Sloan Digital Sky Survey (SDSS; Gunn et al. 2006) uses a 2.5m telescope with a multi-band CCD camera and two spectrographs operating between wavelengths of $3800-6100$ and $5900-9100$ angstroms. SDSS imaging covers a $3000-10600$ angstrom range in 5 filters (u,g,r,i,z). There have so far been three complete phases in the development of SDSS; the first of these, SDSS-I (2000-2005) contains optical imaging over 8000 square degrees, and spectra from selected galaxies within that footprint. The combination of SDSS-I with the second phase of operations (SDSS-II, 2005-2008) forms the SDSS legacy survey expanding the original observations to cover 8,400 square degrees. This includes over 10^8 individual objects and more than 10^6 spectra. The data from the SDSS legacy survey is publicly available in the seventh Data Release of SDSS (DR7; Abazajian et al. 2009).

SDSS-III (2008-2014) has built on the legacy of the prior phases using a major instrumentation upgrade to conduct four simultaneous surveys; this has delivered data products for a wide range of complementary science goals. The SDSS Baryon Oscillation Spectroscopic Survey (BOSS) (Dawson et al., 2013) is the largest of the four surveys in SDSS-III. BOSS was primarily designed to undertake large scale redshift surveys in order to accurately constrain cosmic distances using the BAO method, details of which can be found in Weinberg

et al. (2013). Using a plug plate and optical fibre design the BOSS instrument is able to take 1000 spectra in a single observation. It uses two double spectrographs with blue and red cameras covering a total wavelength range of $3600-10400$ angstroms.

Over its 5 year survey lifetime BOSS has observed 1.5 million luminous galaxies up to $z = 0.7$ and 150,000 quasars in a range of $2.15 \leq z \leq 3.5$; this clearly has impact for many science cases beyond the BAO KSP. In order to benefit as wide a range of goals as possible, and maximise the efficiency of observations, around 12 percent of all targets were given to ancillary projects which had aims not covered by the main survey targets. Ancillary programs were scheduled either in parallel with the core survey targets or as dedicated observations after the main survey was complete. I facilitated the application for a follow-up survey of new LOFAR data using this ancillary time on BOSS. This application was successful and resulted in the project being awarded four hours of dedicated observing time, which we chose to split over four plates giving 4000 potential spectra.

The efficiency of the BOSS spectrographs rapidly falls away at magnitudes fainter than $r = 22-23$. This is shown in Fig. 3.2 from Alam et al. (2015) which shows the r-band magnitude distribution of successful (ZWARNING=0) BOSS redshifts in the most recent data release. As such I use a flux limit of $r_{model} = 22.4$ on our main target list and also include a small number of objects up to $r \leq 23$ at a lower priority in the hope of probing the fainter radio-optical population and filling remaining fibres.

3.2.2 SDSS DR9 matching and target selection

The Ninth SDSS Data Release (DR9) Ahn et al. (2012) was the latest photometric catalogue available from SDSS-III at the time of the radio observations. DR9 is a catalogue of approximately 10^9 objects over an area of $14,555$ square degrees and includes improved astrometry over the previous data release. In order to provide targets for the spectroscopic follow up, radio sources must be matched to known photometric objects in the Sloan database. The next sections detail my method for achieving this.

Approximately 300 of the few thousand radio sources in the observations were identified as multiple component (MC) sources. These are objects with extended emission, or where complex structure can be seen, as described below. The remaining objects, which make the majority of the sample were categorised as single component sources. For these, the optical and radio emission is expected to originate from the same, or nearby, area of the galaxy.

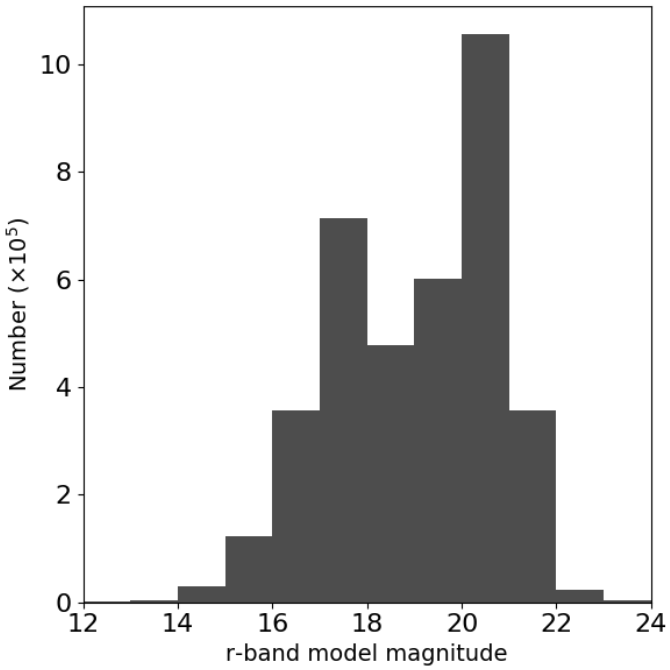


Fig. 3.2 SDSS r-band magnitude distribution of successful BOSS redshifts in the DR12 catalogue.

Hence, any variation in observed position should be small, and mostly due to astrometric error. Conversely, radio emission from MC sources may be up to several arc-minutes from the optical counterpart. In addition to being more difficult to match, MC objects are less likely to have any observed optical counterpart, since they are often faint at optical frequencies. K. McAlpine identified 176 such MC sources in the FIRST data, and I found 112 in LOFAR. Of the 2947 matched objects in the final catalogue, only 93 are MC sources.

Single Sources

A single source radio catalogue from each instrument was matched to the nearest neighbour DR9 optical sources within a radius of twice the typical seeing for SDSS (2”). These matches were then assigned a rank “Gold”, “Silver” or “Bronze” based on their SDSS photometric magnitude. A target was labelled Gold if it was the closest match to that radio source and was brighter than $r_{model} = 22.4$ (2529 sources). Silver sources are all other matches within 2” but also brighter than r_model=22.4 (272 sources). Finally, Bronze targets are the nearest neighbour match, but with $22.4 < r_{model} < 23.0$ (260 sources). Some radio source duplicates existed in this target set due to combining catalogues with different positional errors. These

were removed on the basis of positional coincidence ($< 3''$) keeping only one for the final target list. I performed this matching for the LOFAR data with the remainder split among the team by instrument.

Multiple Sources

SDSS matches to 155 MC radio sources (from the 288 originally detected) were identified. However, only 93 of these were finally observed due to the tiling efficiency.

To select good MC candidates from the LOFAR data I used two methods. Firstly, I identified all components found by PyBDSM that were assigned the same intensity island number. This can be seen in Fig. 3.1 where four PyBDSM detections have been grouped into two sources i237 and i242. Secondly, I used the mismatches between the two catalogues mentioned above. In particular, I identified instances where PyBDSM found multiple component sources within 30 arc-seconds of a single BLOBCAT match, or where there was no unique source found by BLOBCAT within 15 arc-seconds of a PyBDSM match. I then classified each source by eye, comparing cut-outs from LOFAR, FIRST and SDSS, to create a catalogue of likely multiple component radio sources. Finally, I double-checked the LOFAR map by eye for obvious large-scale radio galaxies that were likely to be missed by the aforementioned methods. Two such cases were found: one is centred at RA $243^{\circ}28m15s$ and Declination $+54:15:35$ reproduced in Fig. 3.3; the other is located at RA $239^{\circ}19m58s$ Declination $+54^{\circ}:41':11''$ and has already been spectroscopically observed.

LOFAR MC sources were first matched to DR9 photometric sources within 1 arc-minute. Another team member and I then independently identified the most likely optical candidate. This was done though a combination of over-plotting SDSS positions on the LOFAR map and comparing to cut-outs from SDSS. Only when both agreed on the optical source responsible for the radio emission was the target accepted.

In addition a magnitude cut of $r_{fiber2mag} = 22.4$ was implemented; any targets fainter than this were also rejected. This cut is slightly different from the r_{model} used above and underestimates the brightness particularly at low fluxes. This means some faint sources with $r_{model} < 22.4$ will have been scattered out of our selection by this effect. However, given the small size of this sub-sample, and that the purpose of the flux cut was to maximise the quality of spectra in order to obtain a redshift, the effect is likely to be negligible. In total we

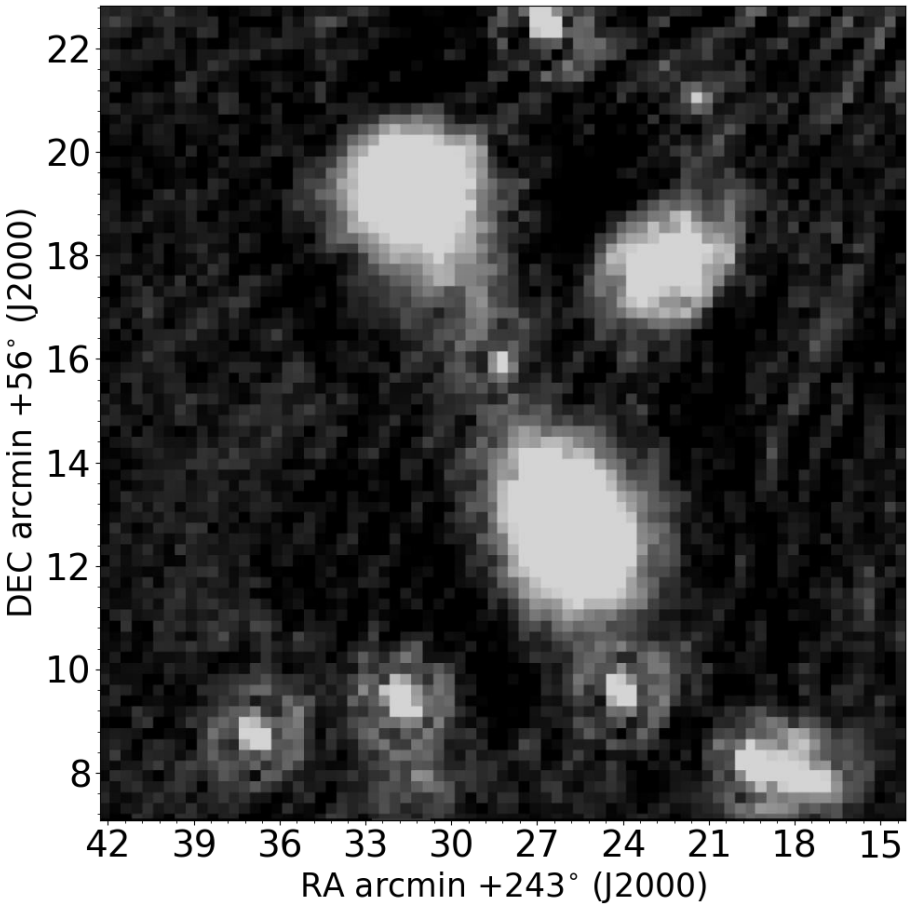


Fig. 3.3 LOFAR ELAIS-N1 cut-out centred at α 243°28m15s, δ +56° : 15′ : 35″ showing a large scale ($\sim$ 10′) Fanaroff and Riley (FR) radio source.

selected 103 targets as matches to MC LOFAR sources, and rejected 9.

Misidentification contamination

There are physical and instrumental variations between observations performed at different frequencies. As such, any match between two catalogues can never be exact. Most objects in this work were paired based on separation distance. This means that some matches may have been made to a random source which was coincidently close. The combination of several radio surveys at different frequencies, with various angular resolution, depth and positional accuracy makes a robust estimate of false matches difficult. However, by comparing our approach to other multi-waveband studies we see that the number of incorrectly paired objects is likely small.

McMahon et al. (2002) matched FIRST sources to a National Geographic Society-Palomar Observatory Sky Survey (POSS) catalogue using a nearest neighbour method. They estimate the number of false matches to a radio source using a random uniform distribution. Within the 2" limit imposed by our matching technique they estimate that 94.5% of the associations are physical.

A similar contamination rate from closest matches between SDSS and FIRST is found by Ivezić et al. (2002) (8% within 3"). However, Best et al. (2005) find that by limited their sample to sources detected above $5mJy$ in the National Radio Astronomy Observatories (NRAO) VLA Sky Survey (NVSS), this contamination falls to less than 1%. Many, although not all, of the selected targets in this work were associated with sources in multiple radio surveys. This suggests a low rate of misidentification to single source objects.

The characteristics of the LOFAR observations are unlike the other surveys in this sample. Therefore, these object pairs may affected by coincidental optical sources in a different way. Optical matches were made to LOFAR sources in the XMM and Boötes fields by Tasse et al. (2008); Williams et al. (2018) respectively. Their results are not directly comparable to this work since they use a likelihood ratio to identify pairs (Sutherland & Saunders, 1992). However, evaluating this likelihood for a 2" separation results in a high probability ($\sim 90\%$) of association for all but the faintest objects. Furthermore, both these studies include several pairs separated by more than 10" which are matched with a probability of over 0.7. This suggests that our 2" limit coupled with the necessary magnitude cut gives this sample a low

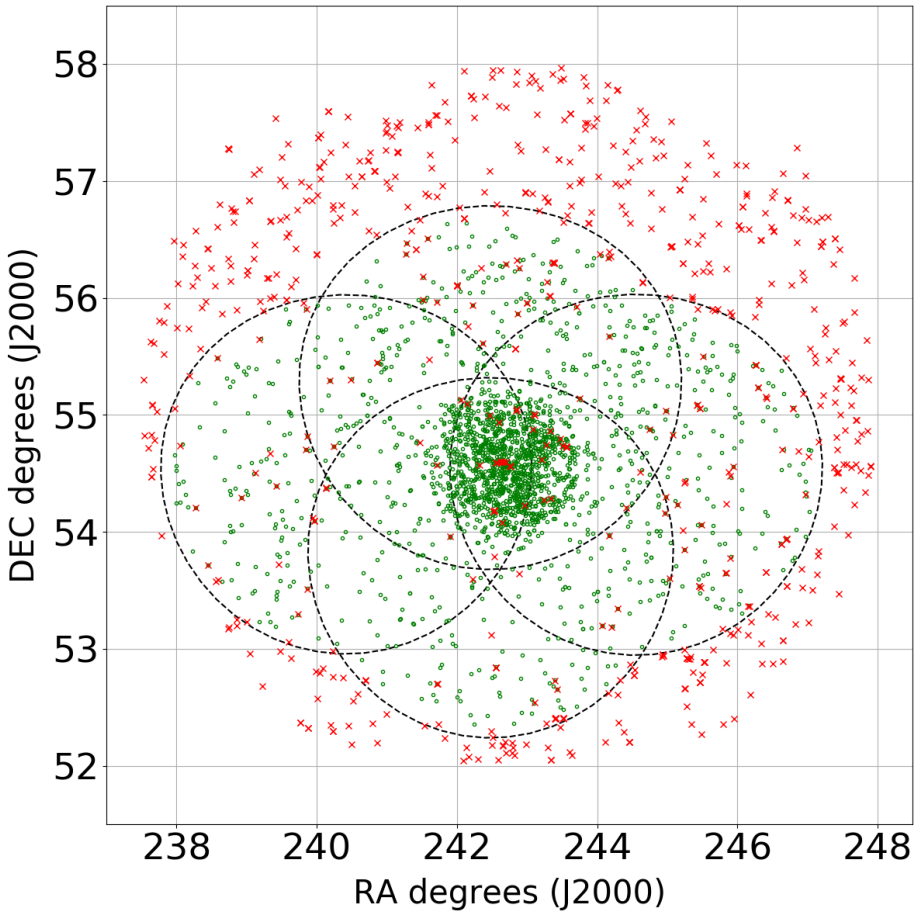


Fig. 3.4 Locations of target sources for BOSS observations. Green circles show SDSS matches to radio targets which were observed, red crosses show location of matches where a fibre could not be placed due to limited tiling efficiency or proximity to another source. Black dashed ellipses show plate coverage. Plate numbers 7562-7565 are shown left, right, down, up respectively.

completeness rather than a high false match rate.

All multiple components matches were performed by hand rather than using separation. The automated cross-matching of multiple component galaxies is a challenging problem. Complex Bayesian fitting methods such as developed by Fan et al. (2015) can outperform humans in some cases, but fall short in others. Hence, there may be some mismatches to these sources. This will make little difference to the overall analysis however since only 34 of the returned redshifts are associated with a multiple component object.

3.2.3 BOSS Observations

Due to the combination of wide LOFAR and FIRST surveys with the deeper JVLA and GMRT data, our target list did not feature uniform sky coverage, with the central area including a high surface density. The restriction on fibre placement (not within 2" of each other), and the desire to observe as many of the LOFAR sources, and hence cover as much sky area as possible, meant that our fibre efficiency was low and only 2194 of 2948 targets could be observed. Fig 3.4 shows the on sky distribution of these targets. Some BOSS LRGs were included in order to not waste the remaining fibres.

3.2.4 Data reduction

The BOSS spectra we originally processed using the main BOSS pipeline, namely the idlspec2d routines, details of which can be found in Dawson et al. (2016). Redshifts and source types were obtained by finding the best fit to the spectra from a set of templates covering the full parameter space expected, given the frequency range of the instrument, and the physical properties of the survey objects (Bolton et al., 2012).

Due to the large number of sources near the centre of the field (Fig. 3.4), two of the four plates (7563 and 7564) lacked sufficient sky fibers and failed processing in v5_7_1 of the data pipeline. The sky subtraction criteria was then altered, reducing the flexibility of the sky model in order to force the data reduction. This resulted in the sky subtraction being worse than a typical BOSS plate, particularly at the red end of the spectrum. The new analysis with the sky model change is labelled v5_7_2 in the catalogue. Fig. 3.5 shows the redshift distributions for plates reduced with different versions of the pipeline. The overall distribution is similar but with v5_7_2 skewed toward a lower redshift, indicating that more distant and hence redder/fainter sources, were more susceptible to decreased S/N.

3.3 Catalogue

The final catalogue is included electronically with the publication Tarr et al. (2017a); I include a sample of the first nine rows in appendix C. The catalogue consists of data drawn from four sources pertaining to the different stages of analysis. I include all 2948 radio sources which were matched to an optical object in DR9 along with selected information from DR9 regrading that object. For the 2194 radio sources that were finally targeted I include the spectroscopic observations; these can also be found in the most recent SDSS data

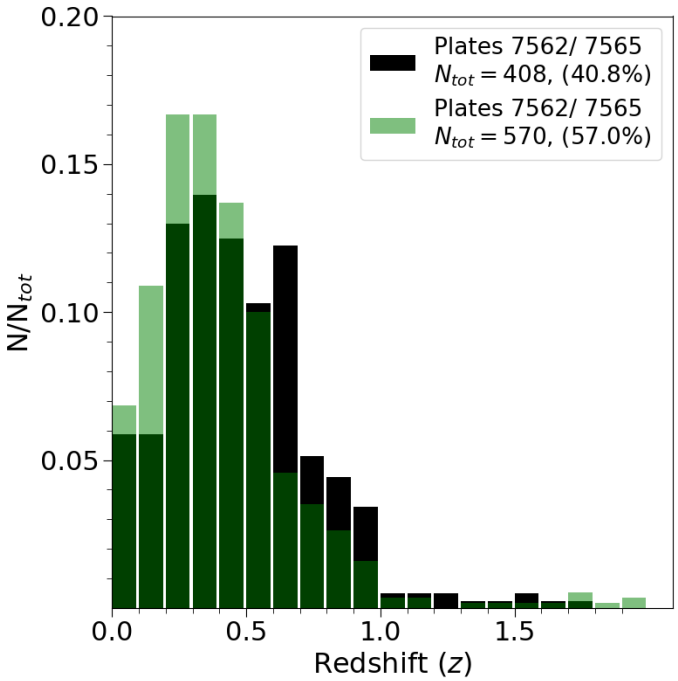


Fig. 3.5 Normalised redshift distribution for objects with Zwarning=0 only. Black solid bars show redshifts from plates reduced using v5_7_1 sky subtraction (7562 and 7565) compared to the re-reduced plates in green. Overall 40.8 percent of spectra reduced using v5_7_1 resulted in a good redshift compared to 57 percent of those reduced with v5_7_2.

release. Finally, I append a visual analysis of the spectra from plates 7562 and 7563 that was conducted by N. Maddox as described in section 3.5.

The following is a description of the important columns for each of the sections above.

Matched radio sources:

SURVEY: A string variable describing which survey the source was detected in.

SURVEY_CODE: Same information as SURVEY, however this is given as an integer, LOFAR=1, FIRST=2 GMRT_garn=3, GMRT_Taylor=4, JVLA=5.

MCLASS: Set to “M” for multiple component radio sources and “S” for single component as deccribed in section 3.2.2

MULTID: Set to same number for assciated sources, set to -99 otherwise.

SAMPLECODE: Quality of match gold, silver, bronze as described in section 3.2.2.

RA_RADIO,DEC_RADIO: Centroid position of radio galaxy emission (or component for multiple type sources)

RADIO_FLUX: Flux in Jy for LOFAR and FIRST sources.

DR9 optical match:

dr9OBJID: Unique identifier for SDSS
modelMag_(u,g,r,i,z): DR9 model magnitude in the five SDSS filters
RA_SDSS,DEC_SDSS: Optical position.
BOSS spectroscopic results:
PLATE: BOSS Plate number 7562-7565
FIBERID: Number of fibre on plate
CLASS: BOSS defined classification from model fit described in section 3.2.4 referred to as CLASS_{BOSS} in section 3.5 either GALAXY, QSO or STAR.
SOURCETYPE: Radio survey identifier, same as SURVEY but for inclusion in main SDSS data release.
z: Best fit redshift, z_{BOSS} in section 3.5.
RCHI2: Reduced chi-squared value of z_{BOSS} to best fit template.
ZWARNING: Quality flag as described in section 3.4.
SN_MEDIAN_ALL: Median S/N of all five SDSS filter bands, used throughout this work, and referred to as S/N.
Visual analysis:
TYPE: Visual object classification referred to in section 3.5 as CLASS_{Vis}, either "g" absorption line galaxy, "e" emission line galaxy, "q" QSO, "s" STAR or "?" unknown.
LINE: Name of line used to identified redshift.
LINEWL: Wavelength of LINE
RESTWL: Rest wavelength of LINE
REDSHIFT: Visual redshift where possible, 0 when no redshift could be determined, referred to as z_{Vis} in section 3.5.
CONFIDENCE: Confidence in REDSHIFT, 1 high confidence, 2 less confident, 3 no confidence or redshift could not be determined.

3.4 Redshift efficiency

A redshift was defined as "good" when its spectrum was assigned no warning flag by the BOSS spectroscopic pipeline i.e. where ZWARNING=0 in the catalogue. For a description of the ZWARNING flag see Adelman-McCarthy et al. (2008). Table (3.3) shows a breakdown of the quality of observations by target type and radio instrument according to the BOSS pipeline. In addition to the number of good redshifts, I also show objects labelled with a ZWARNING of 4 as "poor"; all other spectra are referred to as "bad". Poor fit spectra have a statistically high deviation from the best fit model, but often have high S/N. Galaxies flagged with any other number (bad) usually have multiple issues and are generally unreliable. The

percentage of targets with no quality flag for a typical BOSS survey sample is generally high. Ross et al. (2012) find that 98.2 percent of their high redshift (z>0.42) targets have a ZWARNING of zero rising to 99.6 percent at lower redshifts. This is consistent with the 168 additional LRG targets that were included with this observation, 166 (98.8 percent) of which had no quality flags.

Of the 2175 targets selected for their association with a radio source, 1731 (79.6 percent) were labelled by BOSS as good, lower than for the majority of BOSS observations. Objects classed as a galaxy by the BOSS pipeline had a greater proportion of good redshifts with 87.7 percent labelled as ZWARNING=0 compared to only 46 percent for QSOs. However, the true difference in efficiency may be less than this as (shown in section 3.5.4) galaxies are more likely to be mislabelled (as a QSO) than vice versa, especially those with ZWARNING>0.

Since the key difference between this sample, and the majority of BOSS targets, is the radio selection criteria, it would be reasonable to suggest that the low efficiency seen is dependent on the flux of the radio counterpart. However, Fig. 3.6 shows that the percentage of good redshifts is consistent with approximately 85 percent across the radio flux range observed. There is a slight drop for bright objects (Total flux>0.5Jy) but this is most likely variation due to low numbers in this flux bin (6 objects). This may be a selection effect due to the criterion of $r_{model} < 23$ imposed on the target list. This results in few sources with large radio to r-band flux ratios being observed.

Fig. 3.7 shows efficiency against optical brightness; there is a high proportion of good redshifts up to 20th magnitude in r-band where the success rate drops as expected. Only 55 percent of spectra with modelMag$_r$> 21 have a ZWARNING of zero, dropping to 36 percent above 22. This is lower, but consistent, with the success rate of objects in the ELAIS-N1 region in DR12 at 65.9 percent and 39.8 percent efficiency respectively (Fig. 3.2).

This reduced success can also be seen as a function of S/N as in Fig. 3.8, where the rate decreases to 50 percent at S/N$\approx$ 1. Finally Fig. 3.9 compares the redshift histograms for good spectra against those with a warning label (in red). It would not be informative to plot the success rate as a function of redshift since by definition these “failed” redshifts are uncertain. However, this figure does show that a large proportion of spectra fitted by the BOSS pipeline to a high redshift (z>1) template are found to be problematic. I now consider the quality and accuracy of these fits across the redshift range in more detail in the next section.

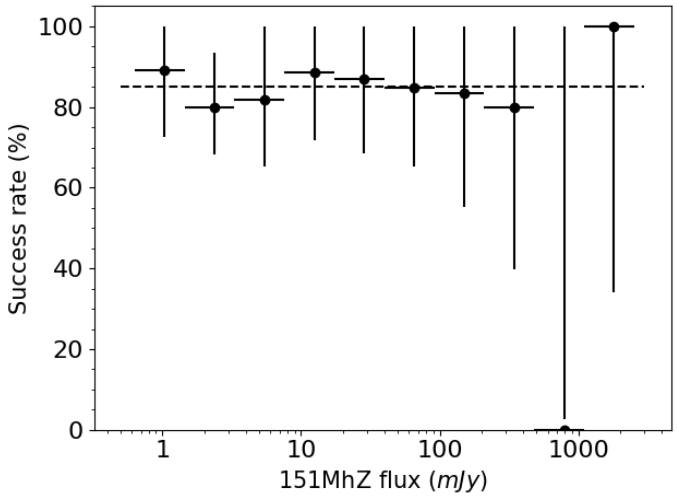


Fig. 3.6 Percentage of "good" (ZWARNING=0) redshifts from LOFAR selected sources in a given 151MHz flux bin. Vertical error bars show 68 percent confidence interval from Poisson noise as defined by Gehrels (1986). Horizontal error bars indicate bin size. Dashed line at 85 percent.

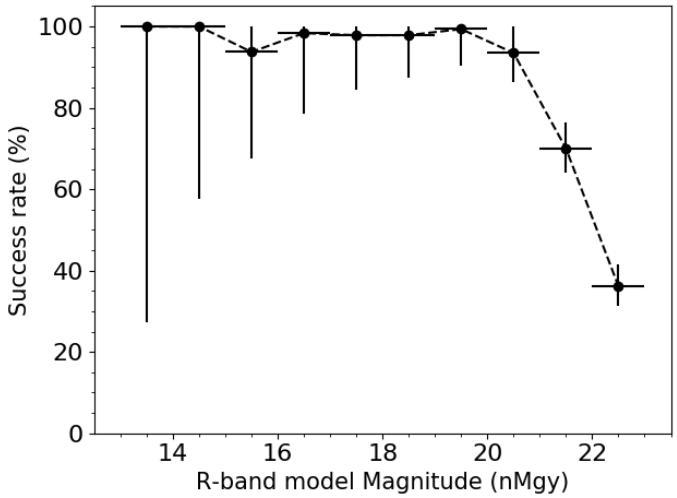


Fig. 3.7 Percentage of "good" (ZWARNING=0) redshifts from all radio selected sources in a given r-band model magnitude bin. Vertical error bars show 68 percent confidence interval from Poisson noise (Gehrels, 1986). Horizontal error bars indicate bin size. Dashed line shows fitted to data points.

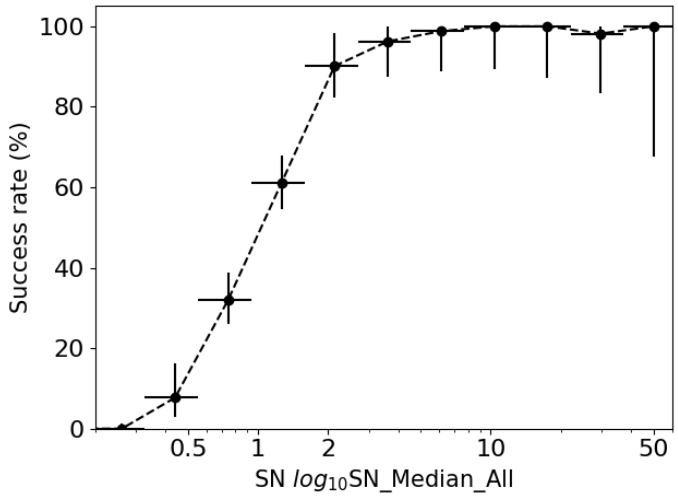


Fig. 3.8 Percentage of 'good' (ZWARNING=0) redshifts from all radio selected sources in a logarithmic S/N bin. Vertical error bars show 68 percent confidence interval from Poisson noise (Gehrels, 1986). Horizontal error bars indicate bin size. Dashed line shows fitted to data points.

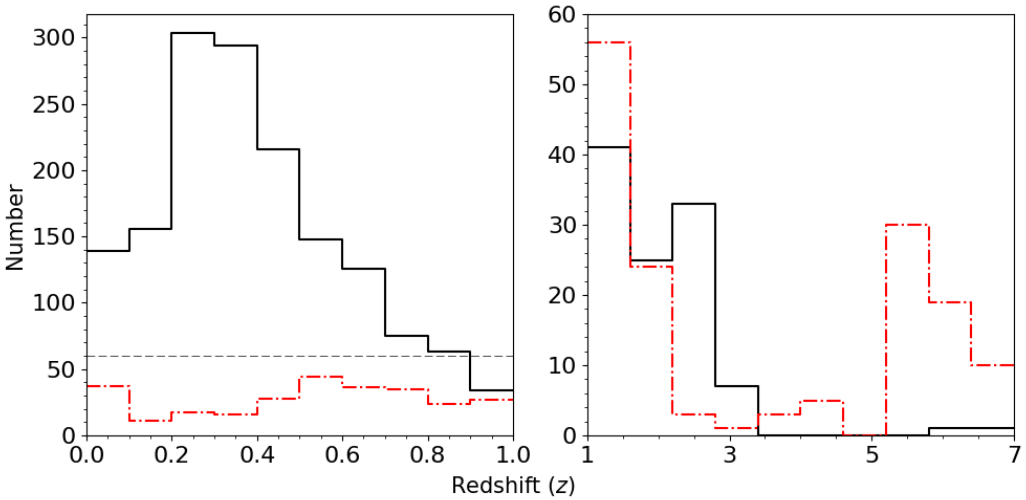


Fig. 3.9 Redshift distribution for all radio selected objects in two redshift ranges. Black solid line shows all spectra with ZWARNING=0. Red dot-dash line shows number count with ZWARNING>0. Black dashed line in left plot corresponds to y-axis limit of high-z distribution.

Source type	Targets	Good	Good Galaxy	Good QSO	Poor	Poor Galaxy	Poor QSO	Bad	Star
LOFAR	411	336	263	33	71	32	30	4	49
FIRST	325	273	224	39	46	25	20	6	11
GMRT	1400	1102	951	104	281	125	135	17	68
JVLA	58	36	29	3	22	9	12	0	5
All RADIO	2194	1731	1467	179	420	191	197	27	133
LRG	168	166	164	1	2	2	0	0	1
SPECTROPHOTO_STD	80	79	-	-	-	-	-	1	79
SKY	1558	-	-	-	-	-	-	1558	-
Total	4000	1992	1631	180	422	193	197	1586	213

Table 3.3 LOFAR-BOSS success rates by instrument and target type. A ‘good’ redshift is defined as any observation with a ZWARNING of 0. Spectra with ZWARNING of 4 have “*More than 5 percent of points above* 5σ *from synthetic spectrum*”, however a reliable redshift can often be obtained where there is sufficient S/N (Bolton et al. (2012)), and are shown here as “poor”. Any observation where ZWARNING is not exactly 0 or 4 is defined as ‘bad’. Note that total columns include a contribution from stars.

3.5 Spectra re-analysis

Menzel et al. (2016) found that using an X-ray selected sample, dissimilar to a typical BOSS selection, resulted in a reduction of the reliability of the redshift fitting algorithm. In particular, errors in redshift and object type were caused by confusion between emission lines, and blending with a neighbour source. From previous surveys, and semi-empirical models, it is expected that this radio selected sample will contain a higher proportion of quasars than the LRG population for which the BOSS pipeline was developed, potentially causing similar issues. The redshift success rates in section (3.4), given by the automatic fitter, show an efficiency for these observations of approximately 80 percent, lower than the spectroscopic success seen in DR10 in general (Pâris et al., 2014).

The BOSS pipeline, or systems based on it, will be used for future spectroscopic follow-up of atypical BOSS objects i.e. WEAVE-LOFAR (Smith et al., 2016). As such, it is important that the performance of the pipeline with this population is investigated. N. Maddox visually inspected half of the spectra (two full plates 7562 and 7563), where possible assigning a visual redshift and type. Since the observations also include standard BOSS LRGs and stars I compare these observations and identify characteristics that correlate with poor accuracy. Additionally, I use the visual analysis to identify the best data from all these observations and combine them to present a preliminary analysis of the radio population selected.

In this work, N. Maddox used code based on RUNZ (Colless et al., 2001) to extract the redshifts from our spectra by visually identifying individual spectral features and calculating the shift required to its rest wavelength. Each visually inspected spectrum in the catalogue includes the spectral future and wavelength used to identify its redshift and class. For a sample of the data and examples of features identified through visual inspection, Fig. 3.10 shows four of the inspected sample.

In total 1239 (50.7 percent) of non-sky fibers were visually inspected, of which 1090(50.1 percent) were of radio counterpart targets. For this section the redshift and source classification (CLASS column in catalogue) given by the BOSS pipeline are referred to as z_{BOSS} and Class_{BOSS}, whereas the new redshift and class from the visual analysis are z_{Vis} and Class_{Vis}.

The visual redshifts were assigned a confidence value (z_{conf}) from 1 to 3. Objects with a confidence of 1 indicate the spectra had clear features and z_{Vis} can be trusted, confidence 2 objects are less clear, often with low S/N, but the observer still thought their identification was correct. Spectra given a confidence of 3 had no emission lines or no obvious spectral

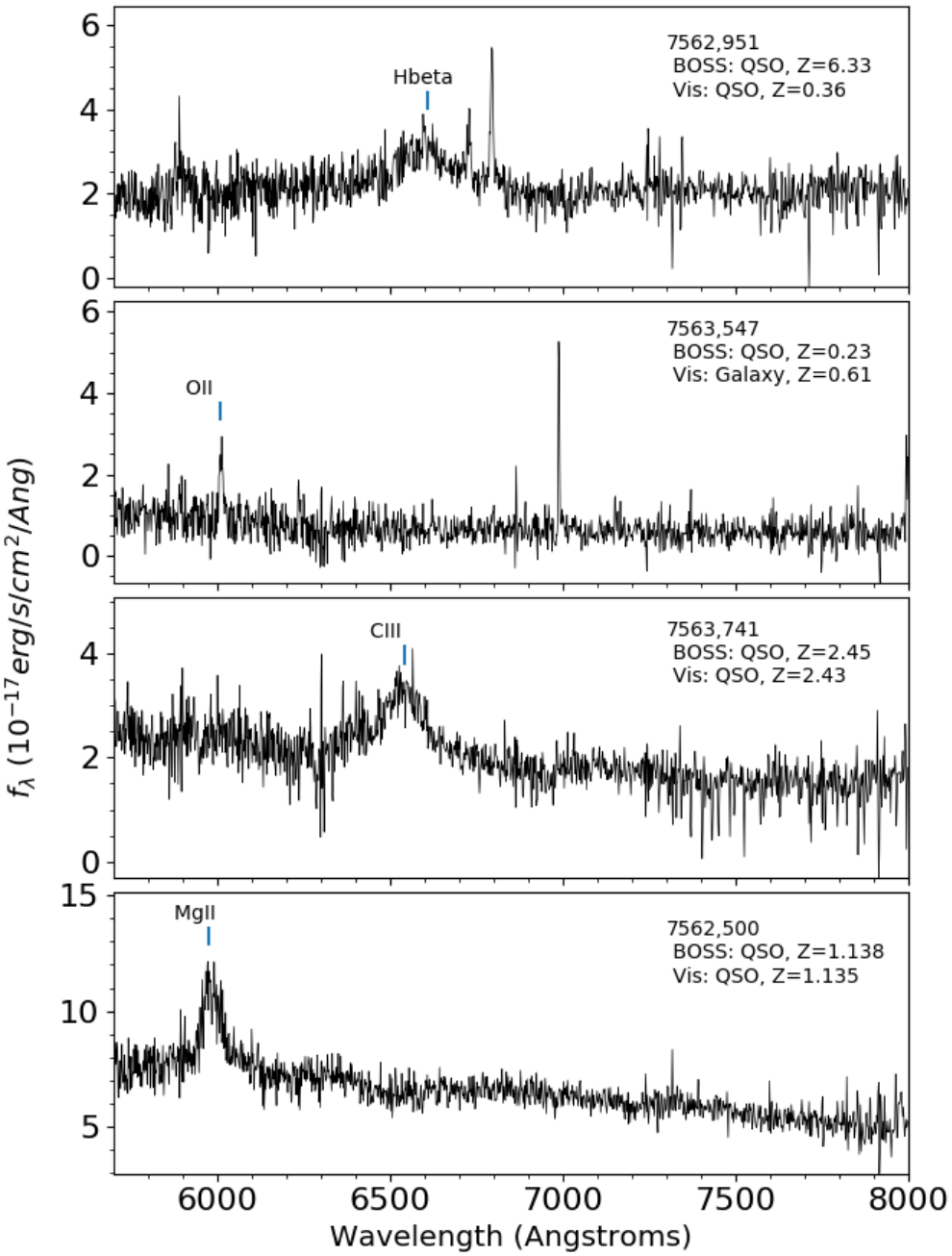


Fig. 3.10 Four of the spectra used for visual redshift inspection showing the spectral feature identified. Top to bottom: 1. Matching class, large discrepancy in redshift; 2. galaxy mistaken by BOSS as QSO, redshift underestimated; 3/4. Visual inspection agrees with BOSS pipeline.

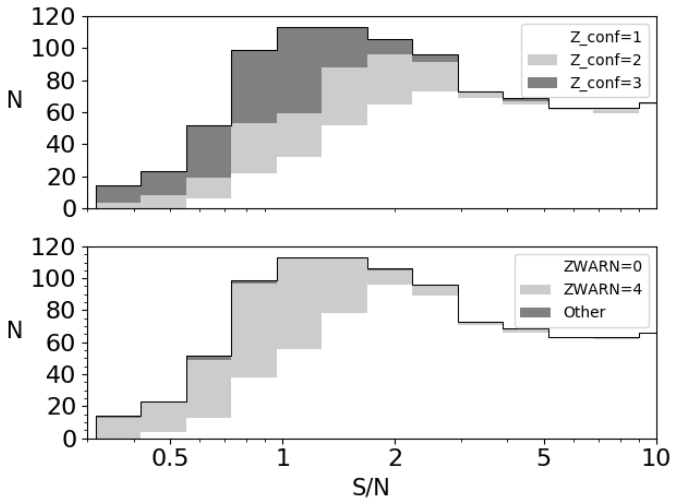


Fig. 3.11 Comparison of visual confidence (top) and BOSS warning label (bottom) of all target spectra in logarithmic S/N bins.

features, and the observer could not produce a reliable redshift (these are assigned z_{Vis}=0).

Regardless of the confidence in the visual redshift, $Class_{Vis}$, “g” (absorption line galaxy), “e” (emission line galaxy), “q” (QSO) and “s” (STAR) can in general be trusted since these are easier to identify than redshifts; cases where the class of an object was ambiguous were labelled “?”. However, objects identified as an emission line galaxy may contain some narrow line QSO’s since these spectra can look similar.

3.5.1 Comparison of visual analysis vs the BOSS pipeline

Fig. 3.11 shows the proportion of high confidence objects increasing with S/N; this trend is comparable to the change in the number of spectra with a ZWARNING of 0. The higher proportion of ZWARNING=0 than $z_{conf} = 1$ spectra at low S/N ($0.5 \leq S/N \leq 3$) suggests that the template fitting method outperforms emission line identification for noisy data, however it could also indicate that BOSS is over-fitting some spectra.

The overall differences between the two methods for both redshift and class can be seen in Fig. 3.13 and 3.14. I started by identifying a threshold to indicate when an error in the BOSS pipeline probably occurred for those objects with a confident visual ID (z_{conf}<3). The threshold for identifying a BOSS error was chosen to be an absolute discrepancy in the redshifts ($\Delta z = |z_{BOSS} - z_{Vis}|$) of at least 0.005. This boundary was chosen such that the 7th to

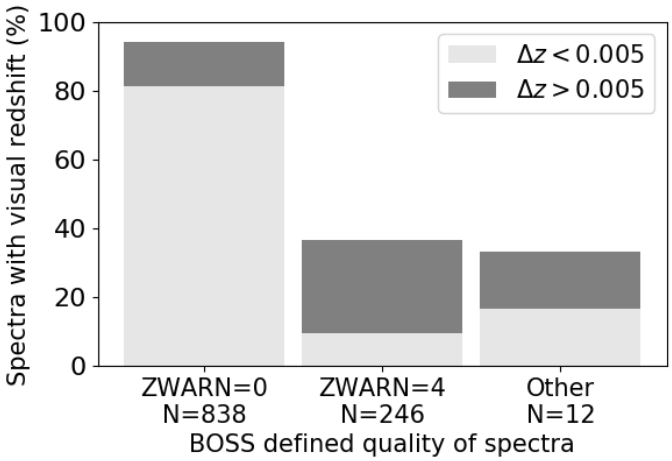


Fig. 3.12 Proportion of objects where the observer had confidence in the visual redshift for each ZWARNING category. Light grey (bottom) indicates a high confidence (z_{conf}=1), darker grey for lower confidence (z_{conf}=2).

93rd percentile of Δz for all z_{conf}<3 objects lie within this range. Hence, the set $\Delta z \leq 0.005$ is considered to be dominated by noise in the visual analysis and unlikely to be an actual error on z_{BOSS}.

Redshift discrepancies

Fig. 3.12 shows the relative percentages of spectra with Δz>0.005 for the three ZWARNING categories. For comparison 143 non-radio selected targets (Boss LRGs and STDs) were also visually checked; only two (1.4 percent) of these have a Δz greater than 0.005 (0.013 and 0.014). The standard deviation of Δz for Boss LRGs and STDs is $\sigma = 0.002$, so in general the visual analysis is robust. As expected the proportion of errors on z_{BOSS} increases for ZWARNING$\neq$ 0. In fact for these cases an error was found with z_{BOSS} over 50 percent of the time when it was possible to discern visual redshift. There are also a small number of cases (6.2 percent) for which an object with ZWARNING=0 also had an error. In section 3.5.2 I investigate the causes of errors on "good" BOSS redshifts and attempt to identify parameters which indicate the highest proportion of redshifts which could be improved through further analysis.

Class discrepancies

I also looked at how the class of objects changed between the two methods. Fig. 3.14 shows the percentage of objects in each ZWARNING category where $Class_{Vis}$ matches,

or is different from Class_{BOSS} and where the visual class was ambiguous. As with Fig. 3.13 there is only a small number of objects (29, 3.5 percent) where a "good" spectrum is given a different class. There are also 42 "good" objects (5 percent) where a Class_{Vis} was not identified. This indicates a noisy spectrum that BOSS was still able to find a model with a good fit to. The proportion of objects with the same class drops drastically however when ZWARNING>0 to only 59 percent of objects where a visual class could be identified. Additionally, the majority of spectra (64 percent) with ZWARNING>0 could not be classified visually anyway. In section (3.5.4) I consider the objects where a difference in class was found in more detail. I then attempt to identify conditions which indicate a higher likelihood of an incorrect class, particularly where further analysis could lead to an improvement.

3.5.2 Redshift error for ZWARNING=0

We next look at spectra where an error in the BOSS fitting algorithm most likely occurred i.e. the difference between z_{BOSS} and z_{Vis} was >0.005. Considering the "high quality" group with ZWARNING=0 and z_{conf}<3, which should have a low number of discordant redshifts, we find there are 49 objects (6 percent of this group) with $\Delta z > 0.005$. Although this is not a large percentage it is significantly higher than the 1.4 percent of LRGs with Δz>0.005 indicating that the BOSS pipeline is less reliable than average for a radio selected sample, even considering the highest quality targets.

However splitting the high quality sample by Class_{BOSS} I find that the majority of objects with $\Delta z > 0.005$ were identified as QSOs, with only 18 (2.7 percent) of the high quality (BOSS) galaxies reaching this threshold. The overall deviation in redshift for these galaxies is also small, with a maximum of $\Delta z = 0.79$. Thus, for spectra identified as galaxies, the pipeline performed with comparable success to typical BOSS observations. Marginal gains in accuracy were far outweighed by the effort of visual inspection.

QSOs

I found that for spectra identified by BOSS as a QSO with no warning label, a Δz of more than 0.005 was present 28 percent of the time (when a visual redshift could be identified). Accuracy for QSOs is important for many cosmological and astrophysical science cases, and the lower number densities of these objects increases the impact any bias will have. As such, it seems prudent to re-analyse all QSO spectra regardless of the quality of the fit as has been previously done with BOSS survey data (Pâris et al., 2017). However, as spectroscopic

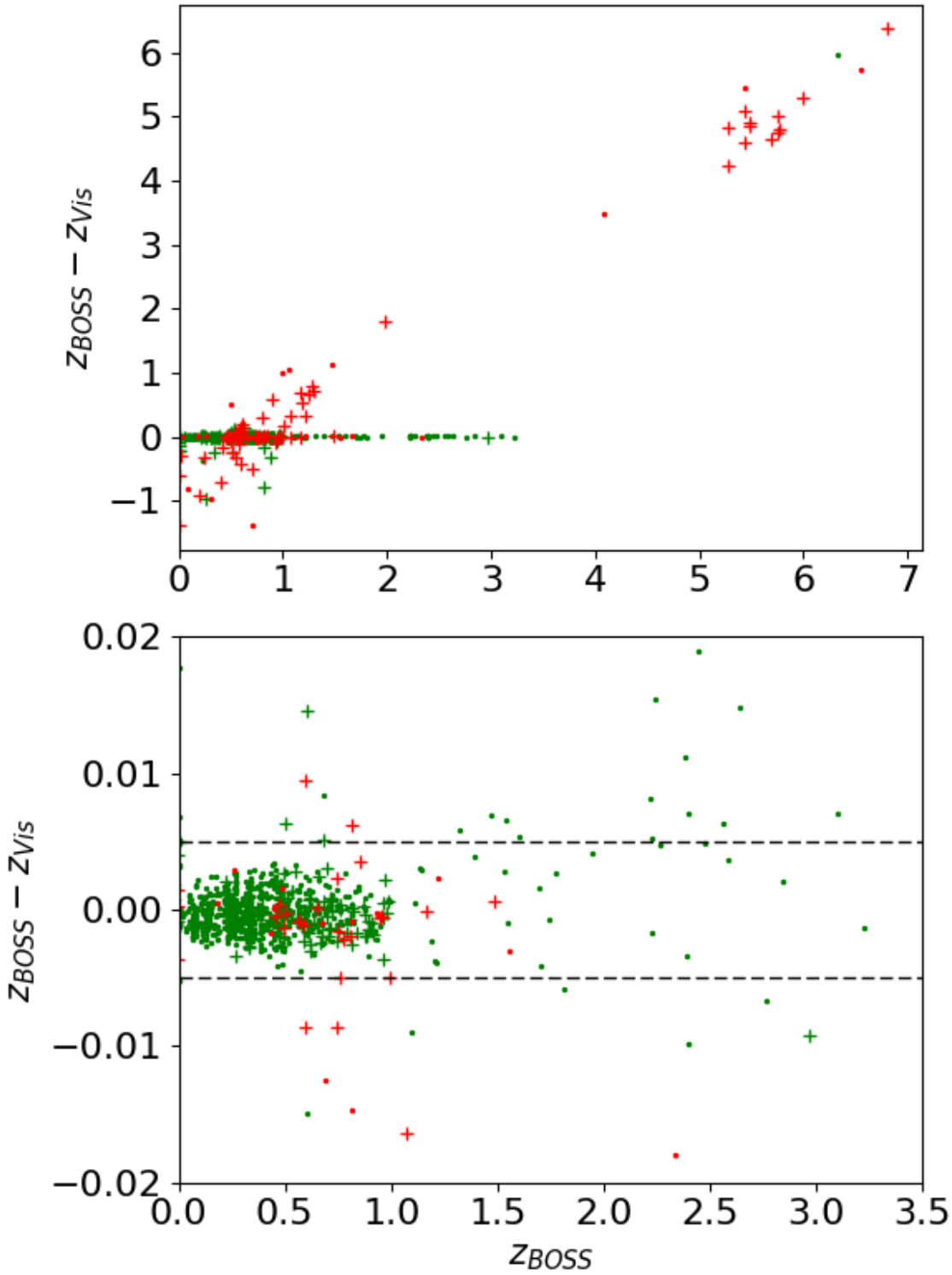


Fig. 3.13 Difference in redshift between z_{BOSS} and z_{Vis} for spectra with z_{conf} of 1 or 2 vs z_{BOSS}. Colour indicates BOSS warning label with ZWARNING=0 spectra in green and red otherwise. Objects marked with "+" have a visual confidence of 2 indicating a larger potential error. Black dashed lines show $\Delta z = 0.005$ threshold used for identifying a likely error on z_{BOSS}.

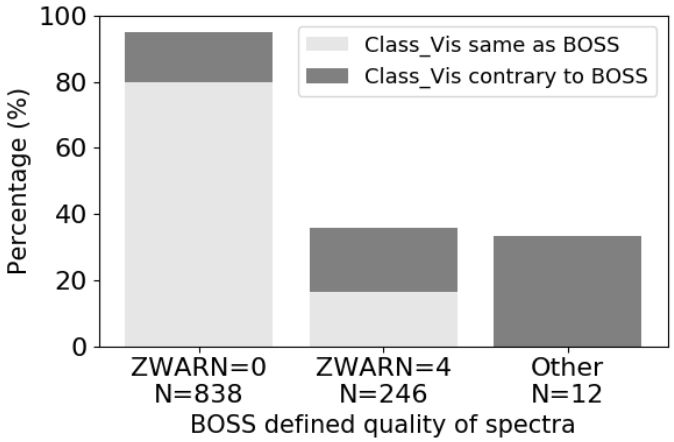


Fig. 3.14 Difference in type classification between identification from BOSS and by eye as a proportion of their ZWARNING category. Light grey shows spectra given the same class by BOSS and visual inspection, objects given a different class shown in dark grey, a visual classification was not possible (Class$_{Vis}$='?') for the remaining spectra.

survey speed increases with the next generation of instruments, this may not be practical and, as shown in Figs. 3.13 and 3.14, large improvements can only be made in a small percentage of cases.

Examining the distribution of QSOs where Δz was large, I attempted to identify regions that contain the highest proportion of errors identifiable through visual inspection. Fig. 3.15 shows the percentage of BOSS QSOs with a ZWARNING of 0 where a visual redshift could be identified (z_{conf}<3) as a function of S/N, r-band magnitude and redshift. The shaded regions indicate the proportion of objects with Δz>0.005.

As expected fainter objects, those at a low S/N and distant galaxies are harder to visually identify. However, this was where there is also the greatest improvement from visual inspection. From considering these spectra the following points are clear about redshifts for QSOs given no warning label from BOSS.

- Below a S/N of 1 few redshifts could be identified with visual inspection; any QSOs at this S/N level should not be trusted but visual inspection is unlikely to help.
- At a S/N level of over 5 the pipeline rarely produces an incorrect redshift; there is little point in re-analysing these.
- QSOs with a model magnitude brighter than 19 in r-band were also accurate.

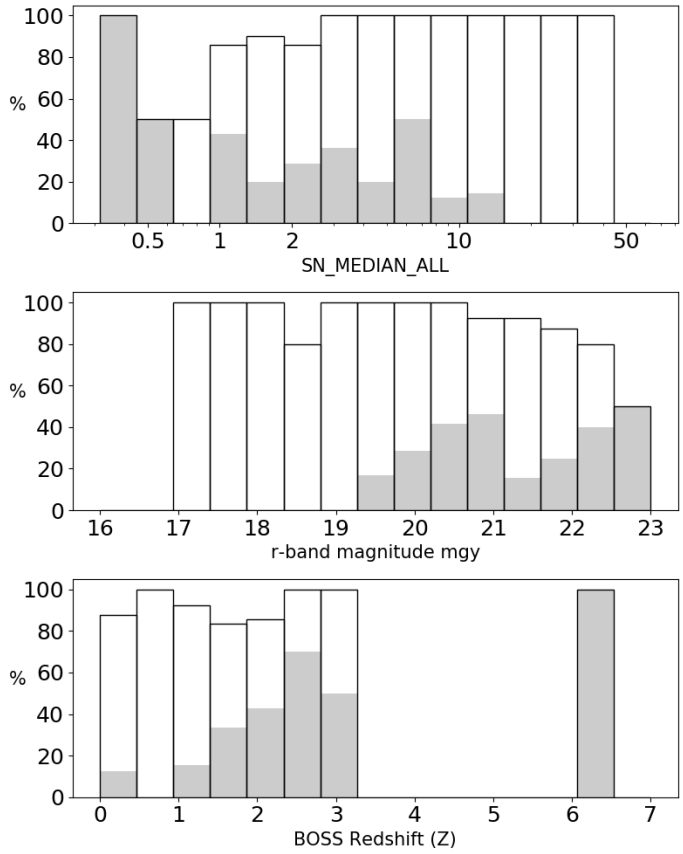


Fig. 3.15 Success rate distribution of visual identification (z_{conf}<3) for radio selected spectra labelled as QSO and ZWARNING=0 by BOSS. Shaded region indicates a probable redshift error (Δz>0.005).

- Faint QSOs with modelMag_r >20 often had redshift errors; at the limit of the instrument modelMag_r >22 visual accuracy also drops.
- At low redshifts (z_{BOSS}<1) identifications from BOSS are generally correct.

I calculated ranges for each of these parameters such that they included 75 percent of all spectra with Δz>0.005 but contained the least number of ZWARNING=0 QSOs overall, maximizing the proportion of improved redshifts. These ranges were z_{BOSS}>1.53, modelMag_r>20.4, and 1.12<S/N<6.6. Clearly these are very specific to this dataset, but similar ranges could be used in future surveys to identify observations that may benefit from further inspection.

For example, ZWARNING=0 QSO spectra with $1 < S/N < 5$, modelMag$_r > 20$, and $z_{BOSS} > 1.5$ includes only 20 percent of that category but over 60 percent of them were improved by visual inspection.

3.5.3 Redshift error for ZWARNING=4

As previously mentioned, spectra with a ZWARNING greater than zero have a much higher proportion of redshifts with a $\Delta z > 0.005$ as seen in Fig. 3.13. However the lower total number of these objects and the decreased confidence (Fig. 3.12) result in only 62 with a confident altered redshift.

Spectra with a ZWARNING not equal to 0 or 4 often have multiple problems beyond a simple poor fit and can in general not be trusted. Even though it was possible to assign a redshift to some of these by eye (33 percent), most were low confidence and almost all differed significantly from BOSS. As such it is not worth re-analysing them and they should be rejected.

Therefore, I now only consider objects with ZWARNING=4. There are 90 such objects (37 percent) where a visual redshift was possible, and of these 59 (66 percent) resulted in an improved redshift. Again there is a higher proportion of QSOs, with Δz>0.005, than galaxies (see table 3.4). However, unlike galaxies with ZWARNING=0, redshifts where there was no confidence in the visual ID cannot be trusted due to the poor fit.

Fig. 3.16 shows the percentage of ZWARNING= 4 spectra where there was confidence in the visual redshift as a function of log S/N for galaxies and QSOs. For QSOs there is no

Class_{BOSS}	Galaxy	QSO	STAR
$z_{conf} = 3$ (No visual redshift)	71	76	9
$\Delta z \leq 0.005$	22	5	4
$\Delta z > 0.005$	19	39	1
Minimum percent with error	17%	33%	7%
Maximum percent with error	46%	89%	20%

Table 3.4 Spectra given ZWARNING= 4 (many outliers) split by Class_{BOSS} and confident ($z_{conf} < 3$) redshift difference. Bottom line shows upper and lower bounds for proportion of those objects with a probable redshift error. i.e the most extreme cases where BOSS redshifts with no visual confirmation are either all correct or incorrect respectively.

S/N limit above which the BOSS redshift is reliable. However, the rate of redshifts obtained from visual analysis does increase over a S/N of 1. BOSS galaxies with a S/N>2 seem to be reliable, and around 40 percent of those with $S/N > 0.5$ were identifiable by eye. Here the visual analysis again reaches its limit at low S/N. There was no confidence in any spectra with a S/N less than 0.4.

There is also a similar trend for QSOs in magnitude as S/N with a smaller proportion of objects given a confident redshift at low brightnesses. To a lesser extent, this same decrease occurs in the quality of visual analysis with magnitude for galaxies. Although the BOSS galaxy redshifts are more accurate than QSOs in general, there is not a brightness limit above which both methods agree completely, or even where the proportion of differences is low. However, the difference in redshift is small ($\Delta z < 0.1$) for bright galaxies ($\text{modelMag}_r < 21$) where a visual ID was possible. These are most likely the same objects with a high S/N.

To maximise the quantity of redshifts and minimise additional analysis and error for ZWARNING = 4 spectra, these data suggest the following. All galaxies above S/N of 2 can be used with no additional inspection, around 40 percent of those above S/N=0.5 and brighter than $\text{modelmag}_r = 22$ will be obtainable by visual analysis, but otherwise are not worth pursuing unless the dataset is small. QSO redshifts in this category are rarely correct and should either all be re-analysed or all rejected. The same S/N and magnitude cuts may be helpful in reducing workload but include the majority of the sample anyway.

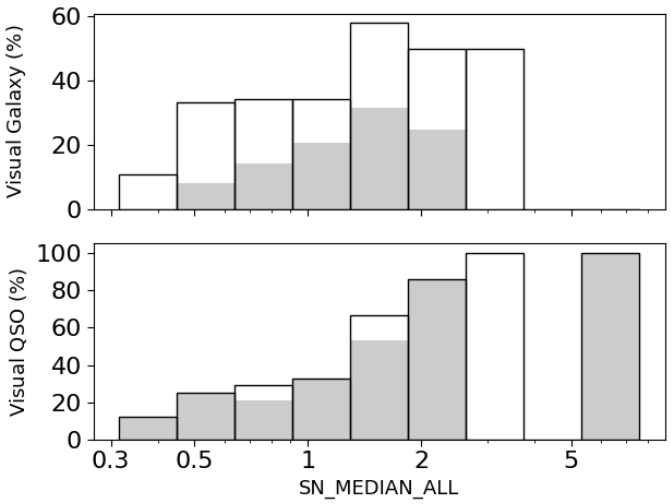


Fig. 3.16 S/N percentage distribution of ZWARNING=4 spectra where a visual redshift was possible (z_{conf}<3), split by $Class_{BOSS}$. Shaded region indicate where $\Delta z > 0.005$. Note low numbers at very large and small S/N; only one QSO and galaxy above S/N=2.75, only 4 QSO and 2 galaxies below S/N=0.33.

3.5.4 Classification error

Fig. 3.17 shows a comparison between the class assigned by BOSS and visually, where an ID was possible ($Class_{Vis} \neq$'?'). Overall BOSS galaxy classifications were correct with only a small number of disagreements. Stars were also assigned the same class visually in almost all cases, although 5 galaxies seem to have been misclassified by BOSS as stars. This seems to be due to many models with a similar reduced χ^2 value confusing the BOSS pipeline. QSO identifications are much worse with around half of BOSS QSOs visually classed as galaxies (where a visual ID was possible).

For BOSS galaxies with a ZWARNING of zero the pipeline does a good job (see Table 3.4). Only 5 (less than 1 percent) of the visually inspected galaxies were mislabelled as QSOs. Although some were not possible to classify by eye, mainly due to poor S/N resulting in a lack of clear single spectral features, the good model fit and low false positive rate otherwise suggests these are trustworthy. ZWARNING=4 BOSS galaxies are less consistent; although there is a similarly low rate identified as QSOs there is a much higher percentage of objects with no visual class at all. These are less likely to be correct since the model fit is worse. However, the proportion of these galaxies which disagree with the visual class, or where a visual class was not possible, is lower when the chi-squared value is close to 1

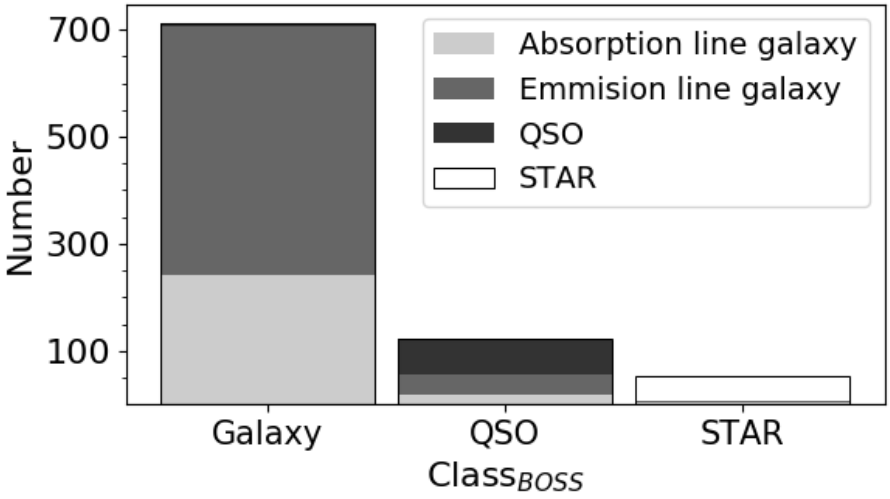


Fig. 3.17 Number counts of spectra where a visual class was possible for each BOSS type separated into visual class category. See Table 3.5 for a more detailed breakdown.

(RCHI2DIFF<0.15) hence these IDs are more likely to be accurate.

Although the BOSS QSO group does contain many objects identified as galaxies by eye, most (92 percent) objects identified as a QSO visually were labelled correctly. The majority of visually identified QSOs which were incorrectly labelled have a χ^2_{red} difference of greater than 0.2. Only 67 percent of BOSS labelled QSOs, with a ZWARNING of 0, are confirmed to be QSOs by visual inspection. There is a connection between the percentage of objects classed differently by the two approaches with both S/N and brightness. 52 percent of "good" QSOs with a S/N<2 are classed differently by eye whereas this is only 17 percent for those with S/N>2. There is also an increase in class discrepancy for fainter BOSS QSOs. Those with modelMag_r>20 were classed visually as an emission line galaxy 30 percent of the time. This decreased to 15 percent for those with a model magnitude brighter than 20 in r-band.

Most BOSS QSOs with ZWARNING=4 are unidentifiable by eye meaning they are unlikely to have been correctly identified given the clear emission lines usually present and the poor model fits. Only 7.5 percent of ZWARNING=4 BOSS QSOs are confirmed by eye; even just considering objects where a visual ID was possible only 21 percent of these classes agree.

	BOSS Galaxies		BOSS QSOs		
Visual class	ZWARN=0	ZWARN≠0	ZWARN=0	ZWARN≠0	Total
Emission line galaxy	435	26	20	18	500
Absorption line galaxy	222	14	0	17	255
Low confidence galaxy	6	2	0	1	9
QSO	5	1	55	11	72
Unidentified	36	75	6	82	208
Total Inspected	704	118	81	129	1090

Table 3.5 Comparison of class identification from BOSS and visual inspection for radio targets.

3.6 Population

I now present a normalised source count as described by Burke & Graham-Smith (2010). I also show the redshift distribution split by object type using the best redshift and class information available given the analysis above. I then compare these to the simulations in Wilman et al. (2008).

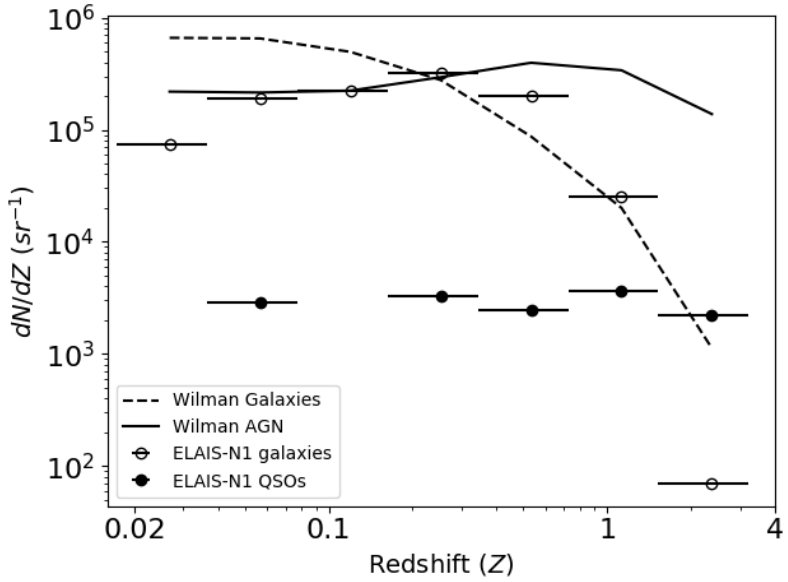


Fig. 3.18 Differential number count of sources normalised by area against log redshift. Points show this work, lines drawn from s-cubed down to $10^{-2}Jy$ at 151MHz. Solid points and line shows QSO/AGN objects, dashed line/ open points show star-forming galaxies.

For the following section redshifts are taken from the visual analysis where available ($z_{conf} < 3$). Otherwise I use the BOSS redshift for BOSS galaxies with ZWARNING=0

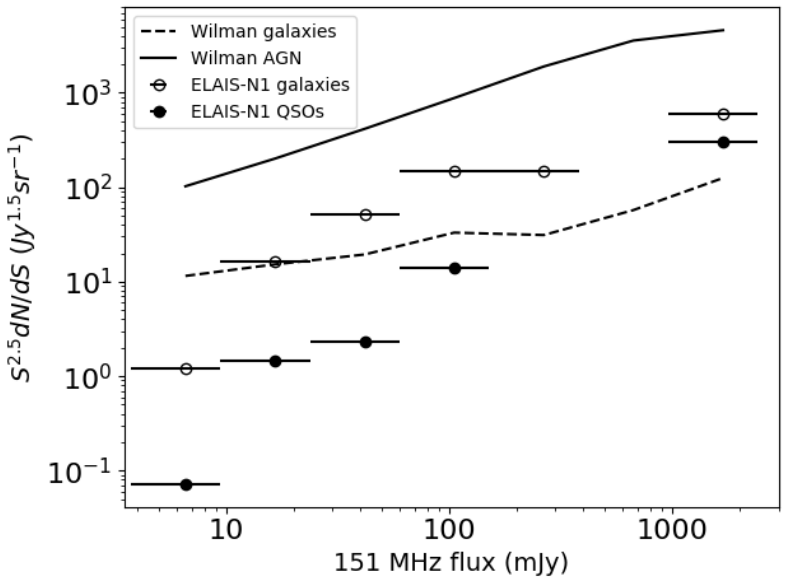


Fig. 3.19 Euclidean-normalised differential source counts for LOFAR 151MHz observation only (points) from 10mJy to 3Jy. Lines for comparison from s-cubed. Solid points and line are QSO/AGN objects, dashed line/ open points show star-forming galaxies.

or ZWARNING=4 and a S/N>2. For BOSS QSOs with no visual redshift, z_{BOSS} is used up to $z = 2$ if ZWARNING=0; for those with a redshift greater than 2, I also impose that $|\chi^2 - 1| < 0.2$ since more than 75 percent of incorrectly identified QSOs were above this threshold. For source classification, I again use Class_{Vis} where possible, reverting to Class_{BOSS} otherwise.

For comparison, I also show the same distributions drawn from the s-cubed simulations down to $S_{0.151} = 10mJy$. The s-cubed model by Wilman et al. (2008) has been shown to follow the radio population found in other surveys closely at these fluxes and wavelengths. As such deviation should be due to a combination of radio survey incompleteness, cross-matching or targeting failure and spectroscopic error.

Fig. 3.18 shows the differential number count of sources as a function of redshift separated by AGN and SF galaxy. For SF galaxies within a redshift range of 0.2 to 2 this work matches the expected number density from Wilman. This is unsurprising since these objects are easy to cross match and the spectroscopic accuracy was very high. At high redshift (z>2) there are many less SF galaxies than expected due to reaching the limits of both radio instruments and the BOSS spectrograph. The low redshift SF galaxies are also under-represented in our data by almost an order of magnitude. This may partly be due to these large foreground

galaxies being resolved out of the LOFAR data or otherwise not being registered by PyBDSM.

For AGN/QSOs there is a good overall agreement in the flat trend of dN/dz although at a reduced factor of about 2 orders of magnitudes. Since this change is independent of redshift it is probably due to a low cross-matching rate and the poor quality of QSO spectra.

The trend in Euclidean-normalised differential source counts is shown in figure 3.19. For AGN we see the broad trend of our data matches that of the Wilman model down to $\sim$50mJy; again the AGN display the lower overall count seen in figure 3.18. Surprisingly however, for SF sources we see an over-representation compared to the Wilman model for fluxes greater than 10mJy. This may be caused by the variance of our sample in this particular field, but will make an interesting issue to study for WEAVE-LOFAR.

A key caveat to the populations presented in this section is that the decomposition by class may not be accurate. Work by Hine & Longair (1979) shows a strong anti-correlation between the luminosity of radio AGN and the fraction with a strong emission line. Thus, it is not expected that the majority of radio AGN will have a strong QSO feature, particularly at these low fluxes. More recently, Kauffmann et al. (2003) find that 8% of the objects classified as galaxies by the SDSS pipeline, are in-fact broad emission line Type 1 AGN. This may, for example, explain the excess number of galaxies seen in figure 3.19. An interesting next step would be to classify these objects using the BPT diagram method developed by Baldwin et al. (1981).

3.7 Conclusions

In this chapter, I have reported on the spectroscopic observations, of a population of faint radio sources in the ELAIS-N1 field. I have presented a catalogue of the matched radio-optical sources from four catalogues and the subsequent spectral observations. The overall success (ZWARNING=0) rate for the spectroscopic follow up was $\sim$ 80 percent although this was much higher ($>$ 95 percent) for bright objects (r_mag$<$ 21). There seems to be no general correlation between the percentage of successful observations and magnitude (except for very faint objects), radio flux or redshift.

In order to test the accuracy of the BOSS pipeline on this atypical sample half the spectra were analysed by eye. Using a single spectral feature, this analysis produced a confident visual redshift and class for 1026 objects. We find a large discrepancy between the BOSS

and visual identifications, especially for objects identified by BOSS as a QSO and those with a poor fit to the model spectra.

Sources classified as a Galaxies with a ZWARN=0 were generally reliable both in redshift and class, i.e. there was no significant deviation between the BOSS pipeline and visual inspection. However, QSOs with ZWARN=0 were classified as galaxies by eye around 50 percent of the time, and there is a significant difference in redshift in 30 percent cases. The redshift accuracy for these objects is worse at low magnitudes ($mag_r > 20$) and high redshift (z>1.5-2). There is also a trend with S/N but it is hard to quantify because of the low confidence in visual redshifts at low S/N.

Objects classed as a ZWARN=4 Galaxy by BOSS were generally confirmed to be such by eye, but where a visual redshift was possible it differed from BOSS around 50 percent of the time; this mostly occurred at low S/N (<1). For QSOs (ZWARN=4) however, the success is much worse. Only 35 percent could be given a class at all by eye and less than 10 percent of the sample were finally confirmed to be correct. For the objects where a visual redshift could be obtained the vast majority (approximately 90 percent) were different for the two methods. There is no S/N, magnitude or redshift cut above which this improved significantly.

The population sampled in this work reflects the expected distributions seen in simulations and other large radio surveys. The efficiency of selecting and observing QSO sources was in general very low. Any future survey would need to improve this as they are an important sub-sample for many science cases including galaxy formation and evolution. Star forming galaxies were under-represented at low flux-densities but the increased sensitivity of recent LOFAR observations and future instruments will improve this. There is also a dip in SF number counts at low redshifts, perhaps due to LOFAR resolving out large low redshift objects.

This project is a good experience step towards the WEAVE-LOFAR survey and other next generation multi-wavelength projects. It has identified areas in which the current fitting technique fails to robustly analyse this data and hence, requires further improvement.

Chapter 4

Gravitational Lensing

Most of the discoveries and knowledge we have developed about objects in the Universe has come from the direct study of the electromagnetic energy they emit. As we have seen in chapters 2 and 3, astronomers have expanded beyond visible light, collecting and combining information from the entire spectrum. However, some things cannot be measured using light at all. It is not possible, for example, to reliably determine the masses of stars or galaxies since the relationship between mass and the stars' radiation is degenerate with many other parameters such as distance, age and metallicity. It is possible to derive these properties using Spectral Energy Distribution (SED) fitting, but the process is dependent on multiple theoretical templates and population models as seen in Capozzi et al. (2016); Maraston (2005). Additionally, there are components that cannot be directly observed with light at all such as black holes, dark matter and dark energy.

As discussed in 1.3.1, galaxy mass can also be measured through observation of rotation curves (Rubin & Ford, 1970). For galaxies which are supported by pressure rather than rotation the velocity dispersion of the galaxy may be used instead (Gebhardt et al., 2000). Although internal galactic dynamics are not a direct probes of mass, they are independent of the stellar composition and thus require far fewer physical assumptions. In order to measure these dynamics a spectroscopic observation is needed at a number of points to create a spectral map of the galaxy. This requirement limits the measurement of stellar kinematics to galaxies which are large and bright. For example, the Mapping Nearby Galaxies at APO (MaNGA) survey (Bundy et al., 2015) selects galaxies larger than $\sim 10''$ which are brighter than ~ 23 AB mag $arcsec^{-2}$ near their outskirts (~ 2 effective radii).

The path that light travels however is dependent on space-time which is in turn related to local mass and energy as described by General Relativity. In practice, we observe the effect

of massive objects as a bending of nearby light rays, a process called gravitational lensing. Lensing allows us to infer the distribution of mass and energy in the Universe completely independent of astrophysical processes.

There are three main regimes of lensing, strong, weak, and micro. Strong lensing is an effect seen in extreme environments such as around super-clusters; it distorts background galaxies into large arcs, rings or multiple images. We have seen in section 1.4.4 how strong lensing has been used to measure cosmological parameters in addition to the mass distribution of the lens. The magnification effect of strong lensing has also been exploited to reconstruct images of distant galaxies at much higher resolution (Jones et al., 2010).

Whilst strong lensing occurs on large scales creating optical distortions of galaxies around clusters, microlensing operates at the other end of the scale. When a bright transient passes behind a massive compact foreground object the source is magnified, seen as an increase in brightness. This technique allows the study of extremely faint objects, such as planets or dwarf stars, as well as the nature of dark matter on small scales (Bennett et al., 2017; Fedorova et al., 2016).

In this thesis I will focus on the weak lensing regime. This occupies the middle ground in terms of the potential that creates it, and where the distorting effect is mostly linear. Unlike strong and microlensing which require suitable alignment of a source and foreground lens, weak lensing probes the general distribution of matter in the Universe allowing study of large scales. This makes weak lensing a powerful and independent probe of large scale structure and fundamental cosmological parameters.

In this chapter I will review the theory of weak lensing and show the mathematical relationship between mass and distortions to light paths. I will then consider the observable effects of weak lensing and the various approaches that have been used to exploit this. Finally, I will discuss the advantages and challenges that are presented by a new generation of radio telescopes and specifically the predicted impact of the SKA for weak lensing cosmology.

4.1 The Path of Light and Refraction

In order to study the effect of lensing on a light ray, it is common to model its path in three parts, a schematic of which is shown in figure 4.1. First a light ray is emitted from a source and travels through flat space-time toward a lens. As the light approaches the lens

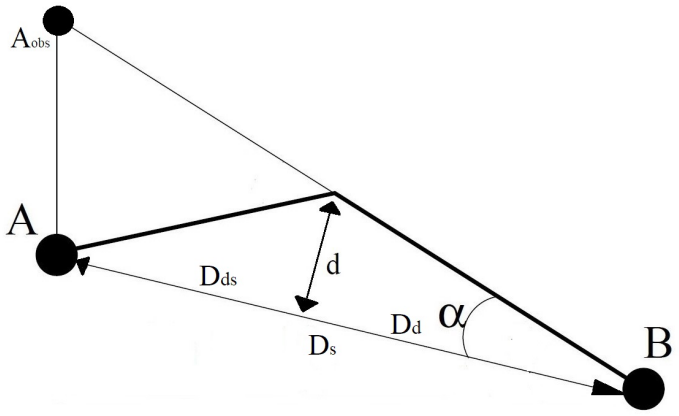


Fig. 4.1 Propagation path of a light ray travelling from a source at A to an observer at B separated by a distance of D_s. The darker path shows the deviation induced by a refractive index between points A and B. This model is useful for describing the effect of a lens on the line of sight.

object it is defected by the curved space-time induced by the mass of the lens. Finally, the light leaves the region of the lens and travels to the observer through unperturbed flat space. For cases where the peculiar velocity and gravitational potential of the lens are non-relativistic ($v << c, \Phi << c,$), the space-time near the lens can be considered to be flat plus the Newtonian potential from the mass distribution of the source, i.e,

$$ds^2 = \left(1 + \frac{1}{c^2}\Phi\right) dt^2 - \left(1 - \frac{1}{c^2}\phi\right) dx^2 \tag{4.1}$$

For the vast majority of astrophysical scenarios, particularly in the weak lensing regime, these conditions hold.

As I will show later in this chapter the bending induced by the lens' gravitational potential can be accurately described as refraction. As such I now discuss the properties of light when travelling through a variable refractive index. For a light ray unencumbered by any space-time curvature the time taken to travel between two points is simply the separation distance l divided by c. Introducing a constant refractive index (n) across the entire intervening medium will reduce the speed of light by a factor of that index giving the travel time as $t = nl/c$. To extend this to describe a variable index we turn to Fermat's Principle of optics,

> "the optical length of the path followed by light between two fixed points, A and B, is an extremum. The optical length is defined as the physical length multiplied by the refractive index of the material".

This implies, as we have already seen, that through a constant refractive index light will travel in a straight line. However, it also reveals that there is an optical length for any index, defined by $n \times l$, over which the speed of light will be c. Hence, we can write the time taken for a light ray to travel between points A and B through a general refractive index $n(x)$ as,

$$t = \frac{1}{c}\int_A^B n(x)dl \tag{4.2}$$

4.1.1 Geometric Time Delay

Let's return to our model of a simple lens in figure 4.1 in which a light ray from A is refracted to B from a point D_{ds} away from A and a distance d from the line of sight between A and B. We can calculate the total travel time of this light using simple Pythagorean geometry as,

$$t = \frac{1}{c}\sqrt{D_{ds}^2 + d^2} + \sqrt{D_s^2 + d^2} \tag{4.3}$$

Multiplying by c provides the optical path length, which we can approximate as,

$$ct \simeq D_{ds} + D_d + \frac{d^2}{2D_{ds}} + \frac{d^2}{2D_d} \tag{4.4}$$

Using $D_s = D_{ds} + D_d$ and $d = D_d \tan\alpha$ we find,

$$ct \simeq D_s + \frac{D_d^2 \tan^2\alpha}{2}\left(\frac{D_d + D_{ds}}{D_{ds}D_d}\right) \tag{4.5}$$

$$ct \simeq D_s + \frac{\tan^2\alpha D_d D_s}{2D_{ds}} \tag{4.6}$$

From this is we can clearly see there is an extra distance due to the bent path when compared to the shortest path between the two points. This gives a time delay such that the longer deviated path takes less additional time than passing through the refractive medium would incur,

$$\Delta t = \frac{D_d D_s \tan^2\alpha}{2cD_{ds}} \tag{4.7}$$

Note that this expression increases with the ratio of D_d and D_{ds}, so for a given bend angle the lens must be stronger if it is closer to the source and thus weaker when nearer the observer. This is very similar to the type of relationship between source and lens that we expect in a gravitation context.

Returning to equation 4.2, which describes path length in a general one dimensional diffraction field $n(x)$, this can be written more generally as a function of three dimensions,

$$ct = \int n(\mathbf{x})(\nabla^2\mathbf{x})^{1/2} dl \tag{4.8}$$

where,

$$\nabla^2\mathbf{x} = \frac{\partial x}{\partial l}^2 + \frac{\partial y}{\partial l}^2 + \frac{\partial z}{\partial l}^2 \equiv \dot{x}^2 + \dot{y}^2 + \dot{z}^2 \tag{4.9}$$

Integrating by parts we can find the difference in time due to a variation of $\delta\mathbf{x}$ along the path, where $\delta x_i = 0$ at A and B, to be,

$$c\delta t = \int_A^B \left[\frac{\partial n(\nabla^2\mathbf{x})^{1/2}}{\partial x_i}\delta x_i + \frac{\partial n(\nabla^2\mathbf{x})^{1/2}}{\partial \dot{x}_i}\delta \dot{x}_i \right] \tag{4.10}$$

$$c\delta t = \int_A^B \delta x_i \left[\frac{\partial n(\nabla^2\mathbf{x})^{1/2}}{\partial x_i} - \frac{d}{dl}\frac{\partial n(\nabla^2\mathbf{x})^{1/2}}{\partial \dot{x}_i} \right] dl + \left[\frac{\partial n(\nabla^2\mathbf{x})^{1/2}}{\partial \dot{x}_i}\delta \dot{x}_i \right]_A^B \tag{4.11}$$

Using Fermat's principle the time between two points must not change and therefore $\delta t = 0$ and δx_i can be ignored. Since we defined δx_i to be zero at A and B the second term is also zero. This results in Euler's equations for the path of light rays,

$$\frac{\partial n(\nabla^2\mathbf{x})^{1/2}}{\partial x_i} - \frac{d}{dl}\frac{\partial n(\nabla^2\mathbf{x})^{(}1/2)}{\partial \dot{x}_i} = 0 \tag{4.12}$$

The unit vector (t) defined by $\frac{\partial x_i}{\partial l}$ will point in the direction of the light ray as it travels through the refractive medium. This vector can be seen as the arrows normal to the light ray in figure 4.2. Substituting this vector into Euler's equations we can reach the expression,

$$\frac{dt}{dl} = \frac{\nabla n - t(t \cdot \nabla n)}{n} \tag{4.13}$$

Since $t(t \cdot \nabla n)$ is the component of ∇n parallel to t, this shows that dt/dl, the change in light path direction, is just the gradient of n perpendicular to t ($\nabla_\perp n$). For the total deflection angle (α) we can integrate dt/dl along the light path as shown in figure 4.2. This gives,

$$\alpha = \int \frac{dt}{dl} dl = \int \frac{\nabla_\perp n}{n} dl \tag{4.14}$$

Hence, the total bend angle of a lensed light path is dependent only on the refractive index field through which it travels. In the next section I will show how the bend can also be

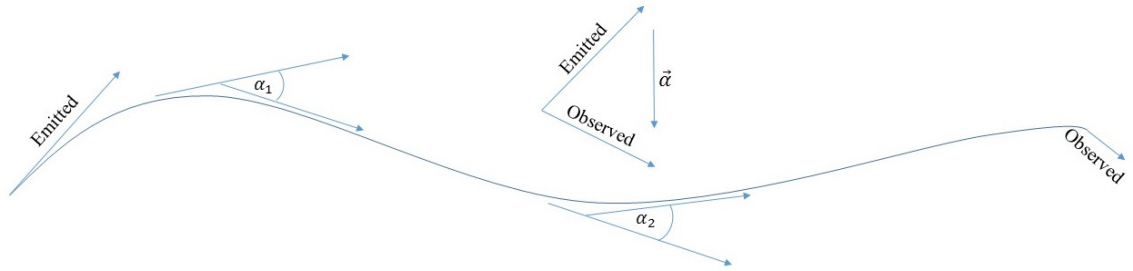


Fig. 4.2 The net bend angle α is equal to the sum of bend angles along its path or, if these angles are small, the difference between the direction of emission and observation.

expressed as a function of gravitational potential. This will demonstrate how astrophysical lensing is similar to diffraction in lab optics.

4.1.2 Relationship to General Relativity

Since $g_{\mu\nu}$ defines the curvature of a particular manifold, it also provides the geodesics that light will travel along when moving through it (Cheng, 2005). A description of the metric near a light ray moving through a potential well, with strength $\Phi(x)$, can be found for the case where $g_{\mu\nu}$ describes only slight curvature. Additionally, it is assumed that the masses involved and reference frame are stationary, and gravitational potential has the usual Newtonian definition $\Phi = -GM/d$. Then,

$$ds^2 = \left(1 + \frac{\Phi}{c^2}\right) c^2 dt^2 - \left(1 - \frac{\Phi}{c^2}\right) dl^2 \tag{4.15}$$

Since light rays trace null geodesics i.e. $ds^2 = 0$, this can be simplified to,

$$ct \simeq \int \left[1 - \frac{2\Phi}{c^2}\right] dl \tag{4.16}$$

If the light is travelling in a flat space, but through a medium with refractive index $n(l)$ dependent on position, using Fermat's principle (Weinstock, 2012) the path length of this light between two points will be,

$$ct = \int n dl. \tag{4.17}$$

Hence, it can be seen that a manifold, with curvature due to a varying potential Φ, can be well approximated as a refractive medium with index $n = 1 - \frac{2\Phi}{c^2}$. Since Newtonian potential is negative for an over-density, this quantity will be equal or greater than zero. Thus we get

the expected result that light travels slower through a potential well. Using this result we can rewrite equation 4.14, for bend angle though a refractive index, in terms of the potential,

$$\alpha = \int \nabla_{\perp} n dl = \frac{2}{c^2} \int \nabla_{\perp} \Phi dl \tag{4.18}$$

Thus the bend angle is related to the perpendicular gradient of the gravitational potential.

4.2 The Effects of Lensing

4.2.1 Bend Angle

There are many effects which can be observed as a result of lensing. The most obvious of these is the delay in time for the light signal. If the space between a source and an observer was flat, then the time taken for light to reach the observer from a distance D would be simply D/c, but for a non-zero potential we have,

$$t = \frac{1}{c} \int_0^D \left[1 - \frac{2\Phi}{c^2} \right] dl \tag{4.19}$$

$$= \frac{1}{c} \int_0^D dl - \frac{2}{c^2} \int_0^D \Phi dl \tag{4.20}$$

$$= \frac{D}{c} + \frac{2}{c^3} \int_0^D -\Phi dl \tag{4.21}$$

Since Φ is negative for an over-dense region, there is clearly a time delay from traversing a curved path. This delay can be measured if a source with high variability lies behind a strong lens (Tyson et al., 1998).

Another directly apparent effect is the bend angle. For any change in the direction of the beam due to varying Φ, we assign a bend angle α. The sum of all bend angles along the line of sight is the net bend angle $\vec{\alpha}$ (Fig. 4.2); this is a function of the integrated potential only.

The bend angle causes light to be detected in a different direction from the one we would otherwise see, so the source would appear to shift position if the lens were removed. With no lensing effects the observer (B) would see the source (A) at an angle β. However, the lensing effect changes the source's position by α, and is now observed at an angle of θ. This gives the lens equation,

$$\vec{\theta} = \vec{\beta} + \vec{\alpha}. \tag{4.22}$$

4.2.2 The Jacobian of Lensing

Consider a source which is not point like but has extended structure, for example, a distant galaxy. Since all the emission does not originate from the same position, α will not necessarily be the same for each photon, since it depends on the profile of the potential field. This mapping from source to image plane can be described by a transformation matrix $\mathscr{A}$ which transforms unlensed to lensed co-ordinates,

$$\mathscr{A} = \frac{\partial \beta}{\partial \theta} \tag{4.23}$$

Combining this with the lens equation, we find that $\mathscr{A}$ can be written in terms of three effects, κ the convergence, γ_1 and γ_2 which are shear terms.

$$\mathscr{A} = \begin{pmatrix} 1-\kappa-\gamma_1 & -\gamma_2 \\ -\gamma_2 & 1-\kappa+\gamma_1 \end{pmatrix} \tag{4.24}$$

The convergence (κ) isotropically focuses the light rays producing a larger image. Since all light rays can be traced back to the source plane, lensing conserves the surface brightness of a source with the total flux proportional to the solid angle subtended by it. Figure 4.3 shows the effect of κ on a circularly symmetric objects in the first row. It is also possible for κ to be negative producing a de-lensing effect which reduces the apparent size of galaxies.

The shear γ_1 and γ_2 are anisotropic mappings that distort the shape of objects through a stretch, thus changing their ellipticity. Closer to massive lenses there are higher order effects that can produce the arcs and rings seen in the strong regime. However, in weak lensing these effects are too small to reliably detect, so I do not consider them here.

4.2.3 Inferring lens or source information

There are two main sources of information that a lensing survey can provide. Firstly, as we have seen, the distortion to background galaxy shapes is directly related to the environment along the line of sight. Hence, it is a key measure of lens mass profiles, the general matter distribution and the geometry of the Universe. On the other hand the magnifying effects of massive foreground objects can greatly improve the resolution of objects in the early Universe if the distorting effects can be removed. However, this is a problem. Here we have

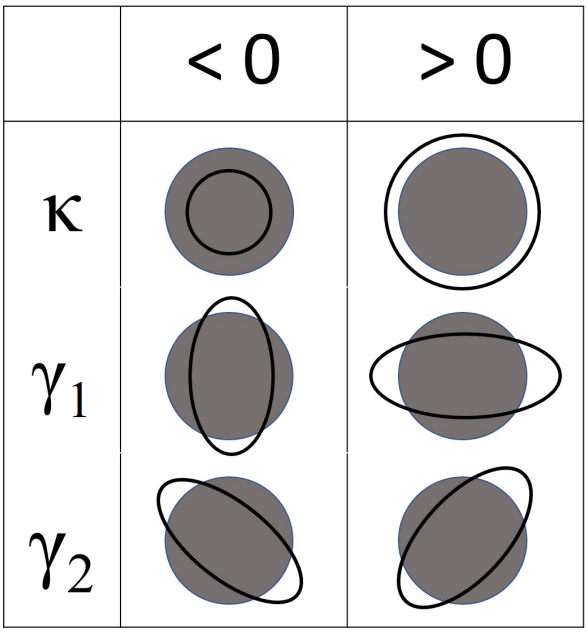


Fig. 4.3 The three second order effects of gravitational lensing.

two unknowns that we would like to explore but only one set of information, the lensed galaxies. We need to know the unlensed shape of objects in order to measure the distortion, or we need to understand the intervening geometry to infer the original shape.

In practice the method for solving this problem varies by the requirements of the project and utilizing other data. For example mass profile models of lens sources can be built based on their astrophysical properties and simulations. Studies of weak lensing generally rely on the stochastic nature of galaxy ellipticities based on studies of the foreground population. I will cover these methods in more detail in the next section. Despite these solutions, the uncertainty in source shape remains the largest source of noise for all lensing projects. However, developments in radio astronomy could provide a new way measure and mitigate this error. This is a potential I explore at the end of this chapter.

4.3 Weak Lensing

In this thesis I will focus on the weak lensing regime where $\kappa, \gamma < 0.1$ i.e the change in a galaxies size or ellipticity is at most 10%. In this regime we can define a "reduced shear" as $\tilde{\gamma} = \gamma/(1-\kappa)$; in this case the lensing Jacobian becomes,

$$\mathscr{A} = (1-\kappa)\begin{pmatrix} 1-\tilde{\gamma}_1 & -\tilde{\gamma}_2 \\ -\tilde{\gamma}_2 & 1+\tilde{\gamma}_1 \end{pmatrix} \tag{4.25}$$

Here κ is only affecting the size of the source. This change cannot be measured directly since it is degenerate with the unlensed size. It is also difficult to construct statistical constraints since sizes vary greatly and are dependent on distance. Changes in shape arise only from the $\tilde{\gamma}$ terms since higher order effects would be negligible in a weak lensing regime. I will use the reduced shear from now on, dropping the tilde.

4.3.1 Measuring ellipticity

In order to extract the components of shear from an image, sources are first identified and their ellipticity measured. However, since these images will be corrupted by several types of noise and resolution limit of the telescope, we require a robust method of making shape measurements.

The traditional technique, still used in various forms, makes use of galaxies' quadrupole moments. If a galaxy has intensity $I(\boldsymbol{\theta})$, where $\boldsymbol{\theta} = (\theta_i, \theta_j)$, then a simple expression of the quadrupole moments is given by Kaiser et al. (1995) as,

$$Q_{ij} = \frac{\int W(\boldsymbol{\theta}) I(\boldsymbol{\theta})(\theta_i - \bar{\theta}_i)(\theta_j - \bar{\theta}_j) d^2\boldsymbol{\theta}}{\int W(\boldsymbol{\theta}) I(\boldsymbol{\theta}) d^2\boldsymbol{\theta}} \tag{4.26}$$

The weighting function $W(\boldsymbol{\theta})$ is used to improve the shape measurements by down-weighting pixels that are expected to be dominated by noise. Typical weighting functions are usually circularly symmetric to avoid introducing additional ellipticity such as a Gaussian or Sinc function. The size of the image will be proportional to the trace $Tr(Q) = \sum_{i=1}^{2} Q_{ii}$ with the off-diagonal elements sensitive to ellipticity. Hence, if the image is of a galaxy with perfect circular symmetry then the quadrupole will take the form $Q = \frac{1}{2} Tr(Q)\delta_{ij}$. The components of ellipticity are then measured as,

$$e_1 = \frac{Q_{11} - Q_{22}}{Q_{11} + Q_{22}}, \quad e_2 = \frac{2Q_{12}}{Q_{11} + Q_{22}} \tag{4.27}$$

This process is not perfect; it also requires precise modelling of the PSF and non-elliptical galaxy morphologies. Over the history of optical lensing there have been many improvements made to the quadrupole moment measurement through further understanding of observational systematics and bias. An alternative method first suggested by Refregier (2003) is to decompose galaxies into orthogonal basis functions called shapelets. The coefficients of

these shapelets can then be summed in different ways to measure the power of various shapes.

Most modern cosmological surveys use a likelihood-based method for estimating the shapes of sources, such as that described by Miller et al. (2013), since they require a higher precision than a quadrupole moment can provide. The latest DES results even use two such Bayesian approaches, *im3shape* and *ngmix* (Sheldon, 2015; Zuntz et al., 2013) to check for consistency. These methods provide a rigorous statistical method for maximising the potential power of a dataset whilst making a priori assumptions explicit. They also have the advantage that uninformative parameters can be properly marginalised over, and can provide a robust measure of noise.

4.3.2 Estimating shear

As discussed in section 4.2.3 the unlensed morphology of an individual galaxy i.e. its intrinsic ellipticity in the source plane (ε_{int}) is unknown. As such, the contribution due to lensing cannot be measured. However, the principle of cosmological isotropy tells us that on large enough scales, there should be no preferred direction in the Universe. Hence, the orientation of galaxies should be random in the source plane[1]. This gives the average of ellipticity of a large sample in the source plane as zero, whereas a small region in the image plane will have a residual signal due to shear. The process of averaging galaxies and the sources of noise which contaminate this measure can be seen in figure 4.4 from the GRavitational lEnsing Accuracy Testing 2008 (GREAT08) Galaxy Challenge (Bridle et al., 2009). The source plane shape of a galaxy as defined by Q can be written in terms of the observed Q and shear by applying the inverse of the lensing matrix $\mathscr{A}$.

$$Q_s = \mathscr{A}^{-1} Q \tag{4.28}$$

Writing ellipticity and shear in their complex forms (ε, γ) and applying the above transform Seitz & Schneider (1995) find the following relationship between source and image ellipticity,

$$\varepsilon_s = \frac{\varepsilon - 2\gamma + \gamma^2 \varepsilon *}{1 - 2Re(\gamma\varepsilon) + |\gamma|^2} \tag{4.29}$$

Where * denotes a complex conjugate and we expect $\langle \varepsilon_s \rangle \simeq 0$. Evaluating this expression to first order Seitz & Schneider (1995) find that,

[1] In fact there are some deviations from this assumption in the real Universe, which I discuss in section 4, but for now let's assume it holds.

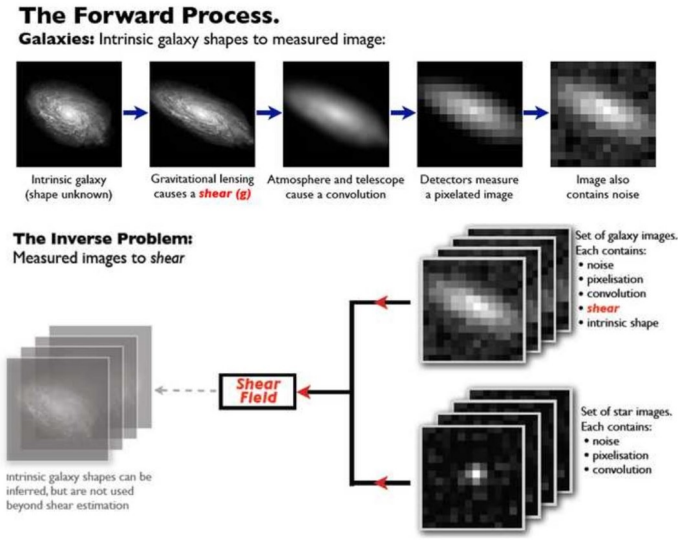


Fig. 4.4 Process of stacking galaxies to reduce ε_{int} CREDIT: The GREAT08 challenge (Bridle et al., 2009)

$$\langle \varepsilon_s \rangle = \langle \varepsilon \rangle - 2\gamma + \gamma \langle |\varepsilon|^2 \rangle \simeq 0 \tag{4.30}$$

providing the shear estimator,

$$\gamma_i \simeq \frac{e_i}{2 - \sigma_\varepsilon^2} \tag{4.31}$$

The term σ_ε^2 is the variance of ellipticity in the source plane. The noise of gamma is therefore dependent on σ_ε and will reduce proportionally with the sample size N as $1/\sqrt{N}$. One could make a shear map by dividing the observation into small regions and applying the shear estimator in each one. The signal to noise of this method is highly dependent on the shear pixel size chosen. Small regions with low source counts will have higher noise. However, on large scales the signal from shear will vary with the expected value tending to zero. As such, a key requirement in lensing surveys is a high density of resolved sources.

4.4 Cosmic shear and mass maps

So far I have described the effect of light rays being distorted in the presence of a single compact object in a lens plane. Measurements that are relevant to cosmology however require knowledge of how the propagation of light is effected by the large scale structure of the Universe. By comparing the measurements of the matter distribution on these scales with

statistics such as the galaxy density two point correlation function we are able to constrain the relationship between mass, energy and structure formation. These are vital components of the cosmological model.

In this section I will present the lensing potential as a probe for a general inhomogeneous Universe. Using this expression, and assuming the weak field regime, I will derive the equations linking this potential, and hence mass, with the lensing effects described previously. Finally, I give the transforms that can be used to construct mass maps from these shear statistics.

Returning to the equation for bend angle in a Newtonian potential (eq. 4.18, γ and κ are known to have the following relationships to ϕ, the 2D projection of the gravitational potential (commonly called the lensing potential),

$$\gamma_1 = \frac{1}{2}(\partial_1^2 - \partial_2^2)\phi, \quad \gamma_2 = \partial_1\partial_2\phi, \quad \kappa = \frac{1}{2}(\partial_1^2 + \partial_2^2)\phi \tag{4.32}$$

The Fourier Transforms of these relationships are,

$$\hat{\gamma}_1 = \frac{1}{2}(u^2 - v^2)\hat{\phi}, \quad \hat{\gamma}_2 = uv\hat{\phi}, \quad \hat{\kappa} = \frac{1}{2}(u^2 + v^2)\phi \tag{4.33}$$

$$\Rightarrow \hat{\kappa} = \frac{|\mathbf{u}|^2\hat{\gamma}_1}{u^2 - v^2} = \frac{|\mathbf{u}|^2\hat{\gamma}_2}{2uv} \tag{4.34}$$

Making a linear combination of these two identities and minimising $\langle\hat{\kappa}\rangle^2$ Kaiser et al. (1995) obtains a best estimator of $\hat{\kappa}$ to be,

$$\hat{\kappa} = \frac{1}{\mathbf{u}^2}\left[(\mathbf{u}_1^2 - \mathbf{u}_2^2)\hat{\gamma}_1 + 2\mathbf{u}_1^2\mathbf{u}_2^2\hat{\gamma}_2\right] \tag{4.35}$$

Hence, we can make pixel images of our shear estimates, Fourier transform, apply equation 4.35 above, and then invert back from the Fourier domain to obtain projected density maps.

4.5 Shear Correlation Functions

In order to use the shear and convergence to constrain cosmology we need to examine the statistical properties on various scales, i.e. its power spectrum (P_κ). This can then be

compared to what we would expect given a certain cosmological model. The power spectrum of convergence can be calculated from a two point statistic of the field combined with a filter function. I now only consider the shear correlation function, which is used in section 6.1.2.

The shear components I have derived in equation 4.32, γ_1 and γ_2 are defined in a Cartesian coordinate system. However, for a correlation between two galaxies we wish to measure their γ components relative to each other, i.e, in a rotated frame such that for objects separated at an angle ϕ from the x-axis,

$$\gamma_1^r = \gamma_1 \cos 2\phi + \gamma_2 \sin 2\phi \tag{4.36}$$

and,

$$\gamma_2^r = \gamma_2 \cos 2\phi - \gamma_1 \sin 2\phi \tag{4.37}$$

The galaxies shears are rotated as 2ϕ rather than just ϕ since γ is a spin-2 quantity. Now we can calculate the pairwise correlation C_{ii} for these particular galaxies of which there are three representing the autocorrelation for both shear components and the cross-correlation between them. For galaxies "a" and "b" separated by a distance l, these would be,

$$C_{ii}(l) = \gamma_{ia}^r \times \gamma_{ib}^r \tag{4.38}$$

$$C_{ij}(l) = \gamma_{ia}^r \times \gamma_{jb}^r \tag{4.39}$$

$$\tag{4.40}$$

The statistic for the whole field is then the average of all pairs within bins of Δl. Due to the symmetry of the cross-correlation we should find that $C_{ij} = C_{ji}$. The cross-correlation should also be zero, making it a test of residual systematics in the data. Using these functions we can also define the ξ correlation which is related to the convergence power spectrum (P_κ) by Kaiser (1992) as,

$$\xi_\pm = C_{11} \pm C_{22} = \frac{1}{2\pi} \int_0^\infty \ell P_\kappa(\ell) \mathscr{J}_{0,4}(\ell l) d\ell \tag{4.41}$$

4.6 Optical Observation Challenges

There are many complex and precise operations that must be performed on an optical observation to produce useful cosmological results. I have provided details of some of these in the previous sections although there are many alternative methods with their own advantages and

problems. In this section I aim to give a broad overview of the challenges in optical lensing studies and thus provide motivation for extending them to other wavelengths as described in the following sections.

Cosmic shear is a statistical measurement due to the random intrinsic ellipticity that is impossible to isolate and measure for a single galaxy. We know that the noise contained in any local estimator of shear will be proportional to $1/\sqrt{N}$. Hence, it is vital that surveys are deep with high resolving power in order to observe small, faint objects and maximise the number densities. Since the precise shape of sources are required, they must be observed with high fidelity. However, at the limit of an instrument's power many objects are smaller than the PSF or the typical seeing creating large distortions in their shapes. Thus, while small, faint sources may reduce noise from the ellipticity dispersion they must be analysed carefully to ensure they do not introduce a bias. In addition to being deep a lensing survey must also be wide if it is to be useful for cosmology and constrain cosmic structure formation and evolution.

It is not just galaxies at the flux limit of a telescope that can be distorted. For ground based telescopes the atmosphere smears images and can mimic a lensing signal. The optics of the telescope itself also contribute to this convolution effect called the PSF. One method for removing this distortion is to measure stars in the FOV that should be point sources and use their ellipticity to calibrate the galaxy shapes. This process can be seen in figure 4.4.

The next challenge before shapes can be measured is to identify and extract sources to be processed. This seems like a straight forward task and is relatively easy for humans given high S/N images. The vast numbers of objects in a large deep survey, often in excess of 10^6, make working by hand impractical and the process must be automated. There are many effective source extraction programs available, for example, *SExtractor* (Bertin & Arnouts, 1996). However, none are perfect and missed galaxies will reduce the signal to noise, particularly if there is a bias against certain ellipticities or shapes. Additionally, if distorted stars or other point sources are identified as extended objects the signal could be systematically biased.

Measures of ellipticity, such as Q or Bayesian fitting methods, use weighting schemes or models that assume an elliptical galaxy profile. For the majority of cases this assumption is accurate and they are an efficient measure of ε. However, there is a small but significant proportion of galaxies that have unusual morphologies. For these cases, it was unclear if the

elliptical model is optimal. This was a key motivation of the GREAT10 challenge (Kitching et al., 2012), which found that greater accuracy could be achieved though careful shape modelling. This information is included in the analysis of current surveys, such as DES, though a calibration of the shear response to the actual galaxy image (DES Collaboration et al., 2017).

Finally, all the work I have presented so far has been based on the assumption that due to the isotropy of the Universe galaxies will be randomly orientated before they are lensed. This assumption is broken when objects in a particular region become preferentially aligned or anti-aligned; we call this intrinsic alignment. These alignments can skew the lensing correlation function in two main ways. Firstly, the formation of galaxies and their resultant orientation is influenced by the gravitational environment they inhabit. If a group of galaxies form in a nearby region of space, within a large scale gravitational field, there will be a higher probability of alignment. This effect is called the Intrinsic-Intrinsic correlation (II) and produces an excess in the lensing power spectrum on small scales. By excluding nearby pairs in the lensing analysis this contamination can be essentially negated at the cost of a smaller data set. Using a distance measure such as redshift can be beneficial in identifying true pairs in three dimensional space, and minimises the loss of data (Heymans & Heavens, 2003).

The second source of contamination is called Gravitational-Intrinsic correlation (GI) and is more difficult to remove. Foreground objects are preferentially aligned with their local potential. However, this potential is the cause of the lensing shear on the source galaxy, resulting in the systematic anti-alignment of foreground and background objects, GI. Sophisticated nulling techniques have been developed by Joachimi & Schneider (2008) that can construct measures of cosmic shear free from intrinsic alignment. Alternatively, DES Collaboration et al. (2017) models this effect and incorporates it into the cosmic shear fitting method. If left untreated GI will produce an underestimation in the power spectrum from cosmic shear.

From this brief overview it is clear that there are several important and sensitive components in the measurement of the cosmic shear signal. Some of these problems are still sub-dominant to the measurement error of current instruments. However, this will not be the case for the next generation of large scale surveys for which statistical errors are expected to drop by an order of magnitude. The methods used by the lensing community must improve and new techniques be considered if we are to overcome these inherent issues and realise the full potential of future instruments.

4.6.1 Recent results from optical weak lensing

The latest cosmological results using optical weak lensing are from the Dark Energy Survey (DES). DES uses 26×10^6 source galaxies and 6.5×10^5 LRGs to calculate and combine three two-point functions of shear-shear, galaxy-galaxy and shear-galaxy. From this it has measured the ξ correlation function, and hence convergence power spectrum, finding a residual systemic of $\sim 10^{-4}$, figure 4.5. From this, and in combination with *Planck* DES has been able to constrain a wCDM cosmological model finding $\Omega_m = 0.23^{+0.023}_{-0.015}$, $H_0 = 78.5^{+2.3}_{-3.7}$ $\mathrm{km\,s^{-1}\,Mpc^{-1}}$ and the dark energy equation of state to be $w = -1.34^{+0.08}_{-0.15}$, as seen in figure 4.6.

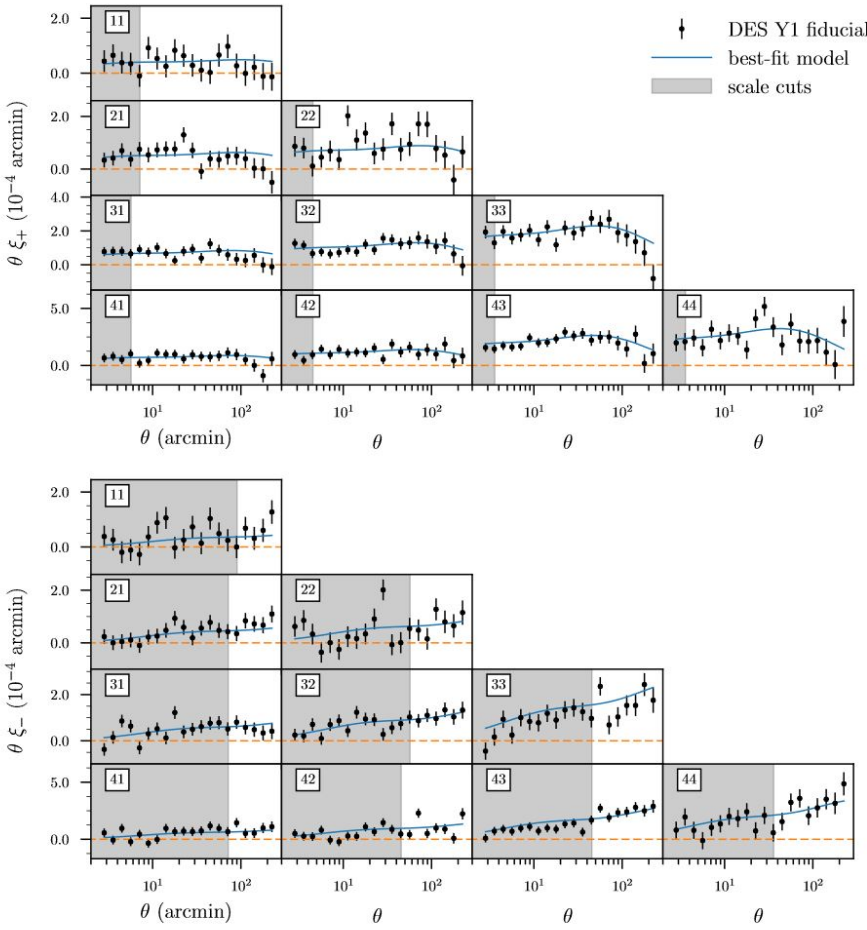


Fig. 4.5 The cosmic shear correlation functions ξ_+ (top panel) and ξ_- (bottom panel) in DES Y1 in four source redshift bins. CREDIT: DES Collaboration et al. (2017)

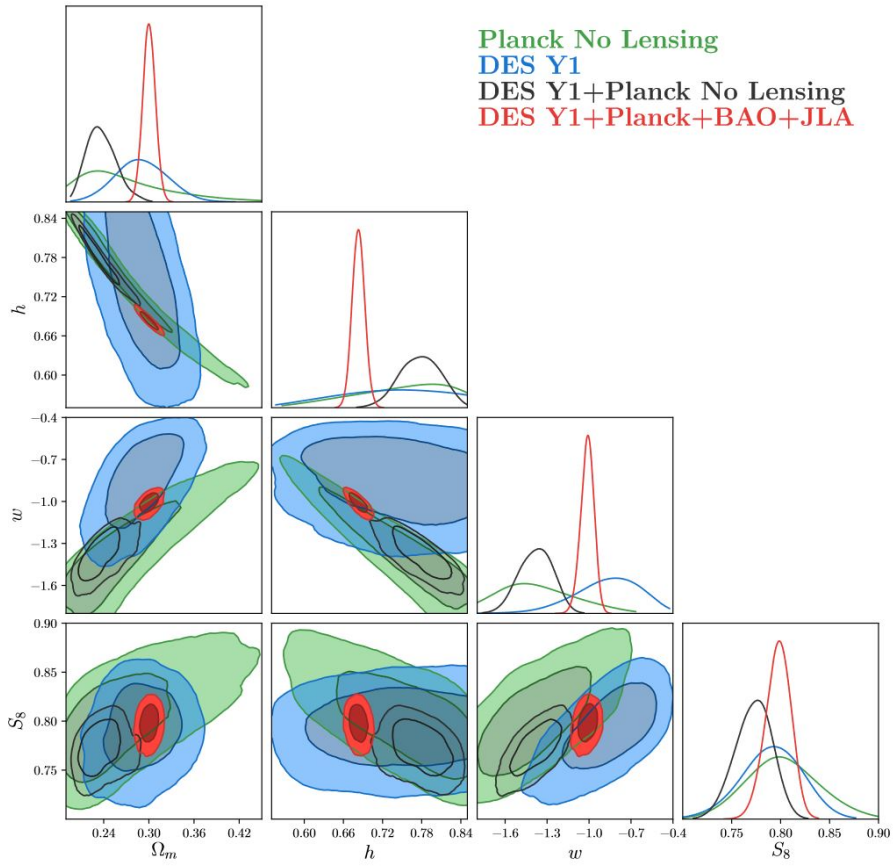


Fig. 4.6 wCDM constraints from the three combined probes in DES Y1 and Planck. Note the strong degeneracy between h and w from *Planck* data. The lowest values of w are associated with very large values of h, which would be excluded if other data sets were included. CREDIT: DES Collaboration et al. (2017)

4.7 Radio Weak Lensing

Until recently weak lensing analysis has only been conducted using observations at optical and IR wavelengths since they provide the deep, high fidelity images that are required. We have seen how these observations face many challenges due to the limited information available. In recent years the sensitivity and resolution of radio telescopes has greatly increased, with next generation instruments expected to be comparable to their optical-IR counterparts. There are several possible advantages of conducting weak lensing at long wavelengths presenting it as an intriguing alternative. Even if traditional methods cannot be beaten by this new source of data, it is still a new probe of lensing with the potential to reduce errors and unknown systematics.

4.7.1 Advantages of weak lensing at radio wavelengths

One of the main sources of noise, aside from ε in an optical lensing survey is the PSF of the instrument. This effect is generally reduced by modelling the distortion and including this in the ellipticity measurement. For radio instruments however, the PSF is defined by the array configuration and antenna response. In principle this is precisely known, highly deterministic for any telescope. We have seen how these PSFs can also be extremely complex. This makes them lengthy to calculate especially for those that vary significantly with frequency. I have also described the difficult and non-linear process in the PSF deconvolution (sec 2.4.2). However, given sufficient computing resources and a robust deconvolution method, we may be able to remove the PSF from radio observations more thoroughly than at shorter wavelengths.

All telescopes, at any frequency, have a diffraction limited maximum resolution given by λ/D. For traditional optical instruments D, the diameter is limited by the ability to physically construct and minutely manipulate large reflectors. For an interferometer, where the signal is combined from multiple receivers, the only restriction is the size of the Earth, or in some cases more, (Hirabayashi et al., 1998). This provides the ability to create radio instruments that have sub-arc-second resolution, and a wide FOV.

As I have discussed in the previous section ε_{int} and IA are a serious cause of noise and systematic in the lensing signal. There are current solutions to these problems but radio observations present another opportunity for constraining the true intrinsic ellipticity. In addition to an intensity map, a radio observation can also produce polarisation information. Brown & Battye (2011) shows that this is correlated with the intrinsic position angle of the

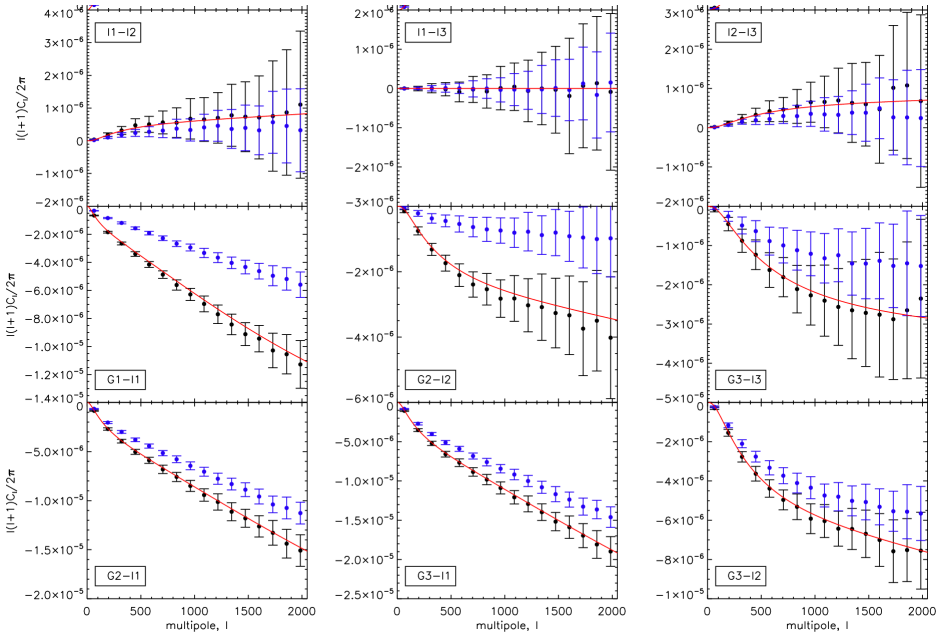


Fig. 4.7 Reconstructions of the lensing and IA auto and cross-power spectra reproduced from Whittaker et al. (2015). In each panel the red curve shows the model power spectra. The black points show the reconstructed power spectra including the intrinsic position angle information from polarization. The blue points show the reconstructions using the standard estimator, as a comparison. From these reconstructions we clearly see that the residual bias has been reduced when including radio polarization.

galaxy. Since polarisation is unaffected by lensing this can be used to reduce noise from shape dispersion and intrinsic alignment. A study by Whittaker et al. (2015) demonstrate that using radio polarization information can significantly reduce bias in the reconstructed power spectra, figure 4.7. In addition Morales (2006) find that HI rotational velocity maps could be an alternative and competitive weak lensing tool completely avoiding ellipticity noise.

There is potential to reduce systemic noise on optical weak lensing using radio surveys even without polarisation or HI velocity maps since radio wavelengths sample a different population. The redshift distribution of sources, at a certain flux limit, tends to skew higher in radio than optical, figure 4.8. This means we can constrain density fluctuations over a greater range of redshift than with optical alone. Also, Harrison et al. (2016); Patel et al. (2010) have showed that there is a dispersion between the intrinsic ellipticities of galaxies in these different wavebands. Thus a cross-correlation would reduce the overall noise from intrinsic ellipticity.

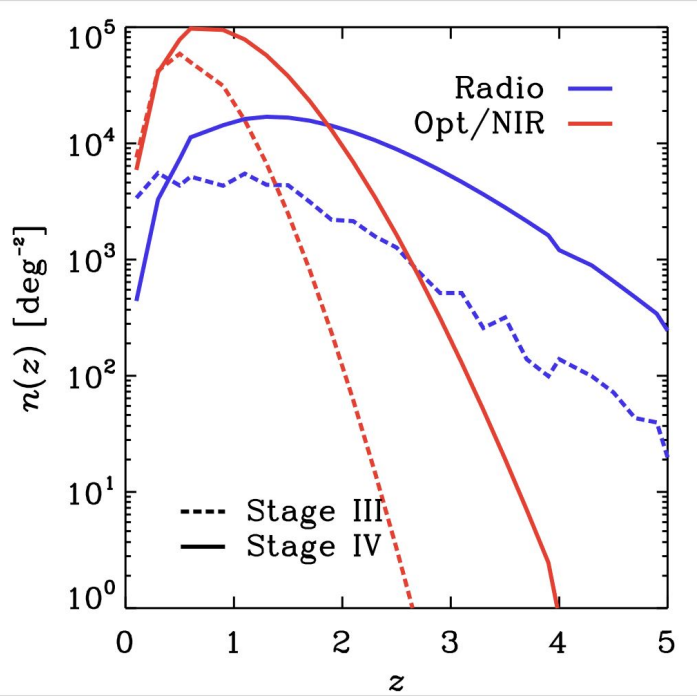


Fig. 4.8 Redshift distribution of sources per square degree. CREDIT: Camera et al. (2017)

Redshift is another key piece of information that can be used to better measure the lensing power spectrum, especially for constraining parameters like the dark energy equation of state. Also, Kitching et al. (2014) have shown specifically that cosmic shear in three dimensions is a sensitive probe of the equation of state. Providing spectra or photometric redshifts for the 10^9 galaxies that next generation instruments such as LSST plan on observing would be extremely expensive. A large 21cm line survey, using SKA2 for example could provide the redshifts required.

As previously stated, low frequency radio telescopes like LOFAR have the capacity to observe wide-field data, from very distant sources, at high resolution. The drawbacks of this ability come in the form of a complex imaging process and vast quantities of data, as discussed in section 2.4. It has been shown however, that weak lensing from radio data is viable, (Chang et al., 2004; Patel et al., 2014). Furthermore, due to the lack of correlation between galaxy morphology in radio and optical frequencies (Patel et al., 2010), a combination of the data could result in a much improved signal. This previous work of weak lensing measurements from radio data is further discussed in section 4.7.3.

4.7.2 Challenges of radio weak lensing

Before the full potential of all these advantages can be realised however, there are some challenges which need to be solved.

Firstly, as discussed in section 2.4, the imaging and deconvolution process involved in making radio images is complex, time consuming, highly non-linear and subject to parameters that can drastically alter image shapes. Recall that a radio image is a model of the data, and it is not always a good one. As such, new bespoke imaging or lensing estimators are required in order to utilise this data. Additionally, these methods must be highly efficient since an observation with the SKA1-mid is expected to produce data volumes of $\sim$ 1TB.

Secondly, the number densities produced by current generation radio telescopes of $\sim 0.1\text{arcmin}^{-2}$ are much less than optical lensing surveys. This will need to increase by at lease one if not two orders of magnitude to be competitive with next-gen optical experiments. Models such as Wilman et al. (2008) and more recently Bonaldi et al. (2016) indicate that the source counts for disk-like galaxies should plateau at fluxes of $\sim 10^{-5}$ and the SKA should achieve the required improvements. However, little is known about the radio population at these extremely low fluxes and it remains to be seen what the true increase will be.

Finally, I have discussed how complex morphologies in optical surveys have required careful fitting methods to achieve the best results. The structures seen at radio wavelengths can be even more unusual with $\sim$ 1 percent of objects displaying multiple lobe or jet components. However, models suggest that at lower flux densities star forming disk galaxies will dominate reducing the relative number of these complex objects.

4.7.3 Previous Radio Weak Lensing Approaches and Results

The first attempt to detect a weak lensing signal at radio wavelengths was presented in Chang et al. (2004), using the VLA's FIRST survey covering 8000 deg^2. Due to the relativity low resolution and number density of sources in the FIRST survey this work considered the lensing signal on large ($1^\circ - 4^\circ$) scales. In order to avoid the complexity and non-linearity introduced by traditional deconvolution and imaging techniques, Chang et al. (2004) estimated source shapes from the visibility data. This method relies on the shapelet approach (Refregier, 2003; Refregier & Bacon, 2003) in which an arbitrary surface brightness profile is linearly decomposed into an orthonormal basis set. Hermite polynomials are a useful choice of shapelet for measuring shape since they represent distorted circular Gaussians. Additionally, the Fourier

transform of these basis functions results only in a scaling and are thus convenient to model in the u, v plane. Chang & Refregier (2002) modelled a VLA observation of these shapelets and fit them to simulated data using a χ^2 minimisation. They found that the recovered shapes from this method were in good agreement with the simulation, and that reconstructed images from FIRST data compared well with those created using CLEAN. From this method of shape measurement Chang et al. (2004) detected an E-mode signal at the 3σ level and no significant B-mode. The initial signal recovered from the FIRST data contained several sources of systematic error, primarily from the anisotropic beam shape of the VLA due to discrete u, v sampling. Despite contributing a smaller effect, all other systematics such as w-projection, averaging, and source confusion had to be carefully controlled in order to prevent them dominating the $\sim 1\%$ weak lensing signal. As these effects are correlated in a complex way they could not be easily removed analyticity. Instead the systematic shear signal was modelled as a function of observational and source parameters and subtracted from the result.

The approach of shape analysis using shapelets was repeated by Patel et al. (2010) using CLEANed images from new deeper VLA data, and a high resolution survey from the Multi-Element Radio-Linked Interferometer Network (MERLIN). This resulted in a much higher number of resolved radio sources over a 70 arcmin^2 region. They also applied the shapelet technique to optical HST data in the region, and used the same shear estimator on both sets. Comparing these two measurements they find a good agreement between the shear estimation distribution from optical and radio. This indicates that the two sources provide similar noise properties in this sample. In addition, the correlation between the shape of matched radio and optical sources was also measured. They found no apparent strong correlation with the Pearson coefficient of < 0.1. This result suggests that a radio-optical cross correlation could help to reduce the shape dispersion noise in shear measurements. This finding does somewhat contradict the results of Battye & Browne (2009) who show that majority of optical-radio pairs in the FIRST/SDSS surveys, are aligned along their major axis. However, these two studies are not directly comparable since one has much higher resolution and the other a larger sample. As such, this discrepancy could be due to a selection effect or statistical uncertainly. Thus, more work is needed to characterise this relationship.

The estimation of shape using the shapelet method on CLEAN images was further tested by Patel et al. (2014) on simulated eMERLIN (data). They found the original ellipticities could be robustly reproduced, finding no additive bias and multiplicative bias of 10%.

More recently Rivi et al. (2016a,b) presented a Bayesian method for fitting ellipticity in the visibility domain called RadioLensFit. This approach uses galaxy models which are defined by an exponential profile distorted by the ellipticity variables e_1 and e_2. It is similar to the shapelet method discussed by Chang & Refregier (2002), since interferometric models are also fitted to the data using a χ^2 method. However, by marginalising over parameters such as brightness and position RadioLensFit produces a likelihood function of ellipticity directly, rather than having to infer them from the shapelet coefficients. Shear can then estimated in the usual way with the variance of the likelihood informing the weighting scheme. For simulated postage stamps of SKA1-MID visibilities RadioLensFit was found to have a biases which would satisfy the requirements of a $5000 deg^2$ SKA1 survey. This technique does rely on the ability to produce visibilities of single galaxies which is currently a complex and non-linear process. As such, further work on source extraction techniques is needed.

4.7.4 SKA Weak Lensing Cosmology

The SKA represents a step change in the power of radio telescopes and looks set to be a premier next generation cosmology tool, competing optical experiments. As we have discussed there are several synergies with optical weak lensing that can improve the measurements. Camera et al. (2017) has shown that by combining optical and radio surveys many systematics in the power spectrum, calibration errors or intrinsic alignment bias can be removed. Figure 4.9 shows this result for the power spectrum systematics. Clearly, SKA cross correlation can provide fantastic improvements for weak lensing cosmology.

In this chapter I have introduced weak lensing as a powerful tool for cosmology. I have also considered the potential of radio weak lensing as a complement to optical studies. I will now attempt to construct a new measurement method for radio weak lensing, which will be of use for current SKA pathfinders and beyond.

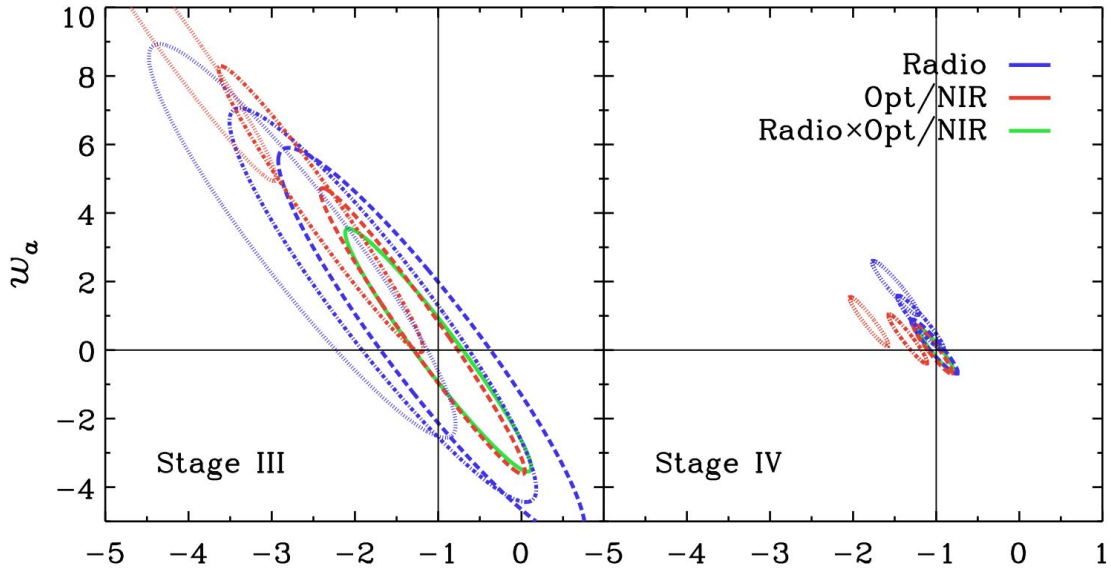


Fig. 4.9 Marginal joint 1σ error contours in the $w_0 - w_a$ parameter plane for stage III and stage IIII cosmic shear surveys. Black cross indicates the ΛCDM fiducial values for dark energy parameters, namely, $w_0 = -1$, $w_a = 0$. Dashed, dot–dashed and dotted contours refer to the residual systematic power spectrum with $\sigma^2 sys = 10^{-7}$, 10^{-6}, 5×10^{-5} respectively. All contours but those for the cross-correlation are biased. CREDIT: Camera et al. (2017)

Chapter 5

Estimating Lensing Shear from Radio Visibilities

The work described in this chapter has been conducted in collaboration with David J. Bacon and is based on Tarr et al. 2017b (in prep).

As we have seen in the previous chapter, radio astronomy can provide a powerful new probe of the gravitational lensing signal, with the SKA set to become a premier cosmology instrument. Previous work has demonstrated that it is possible to detect a shear signal using radio data. Generally these methods measure the ellipticities of individual radio sources, either from the image plane or isolated visibilities, and apply the traditional lensing statistics. In this chapter I investigate the effect of a general lensing field on radio data and hence the possibility of estimating shear statistics directly from the visibilities.

From this analysis I develop a bespoke radio lensing estimator called Fourier Inspection of Lensing Modes (FILM). In section 5.1 I describe the methodology behind FILM and present the two modes in which FILM can operate, pixel reconstruction or Fourier component estimation. In order to test this method I use simulated radio observations based on semi-analytic models[1]. I then present results for various scenarios in section 5.3. Finally section 5.4 provides conclusions and a discussion of the areas in which more work is required.

[1] For details see section 5.2.

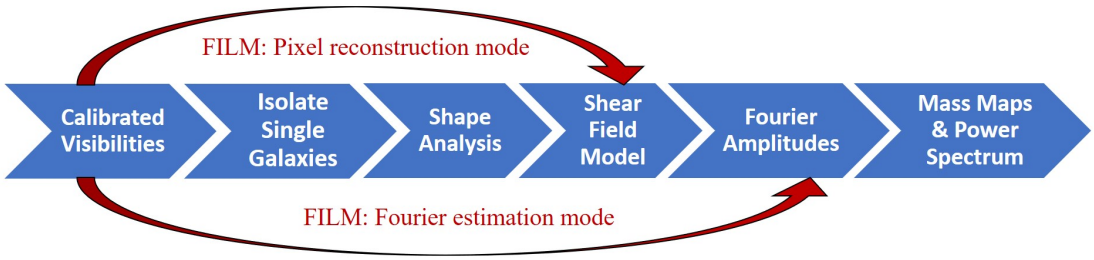


Fig. 5.1 The main steps in producing usable cosmological output from a radio lensing survey with the alternative method "FILM" shown in red.

5.1 My approach: Direct estimation of shear

As seen in section 4.7, when a radio observation contains multiple sources individual galaxies must be isolated before shape measurement can take place. This separation is most easily achieved in the image plane but can also be done through a process such as peeling (Noordam, 2004) where self-calibrations on individual sources can remove most contamination from other objects in the FOV. Figure 5.1 shows the main steps involved in these approaches to radio lensing. As previously discussed the multiple models, transforms and non-linear processes in these methods may make them lengthy and sub-optimal.

The alternative I consider here is a direct estimation of the shear signal from all observed galaxies simultaneously. I have investigated two modes in which a lensing signal could be produced, represented in figure 5.1 by the red arrows. The amplitudes of lensing on different Fourier scales are used to construct the correlation function as described in section 4.5 and are hence very valuable. These are measured by FILM working in Fourier estimation mode. The alternative mode of operation is pixel reconstruction. This aims to produce a map of the aggregated ellipticity signal in some small region of image space that can analysed using standard lensing statistics.

5.1.1 Analytical model of a radio observation

To begin let us consider an analytical model of an idealized radio observation. This will reveal the relationship between a generic shear field $\gamma(\mathbf{x})$ and distortions in the subsequent radio visibilities. I will then use this to form the basis of the estimator in section 5.1.2. In order to maintain the simplicity of the model, and allow insight, I have made some approximations in its construction. To ensure the model is still representative I make a comparison with a numerical simulation finding a negligible difference of the order γ^2.

Sky brightness model

The "true" sky brightness distribution to be observed is a combination of emission of all objects in the FOV. As such, I first define a general source of arbitrary shape and size and sum n such objects for a field of multiple galaxies. Since point sources and diffuse emission provide no lensing signal I do not include them at this stage. However, this emission will be a source of noise in our data and I consider this effect on the estimator later in the process.

Most radio objects can be estimated as an elliptical object or, in the case of multiple component sources, the sum of elliptical components with various profiles. Thus a single galaxy or component in the centre of the source plane with an unlensed emission profile f can be described as $f(\mathbf{x}_s)$, where x and x_s represent image and source plane coordinates respectively. As this source is lensed into the image plane it will undergo a coordinate transformation as described by equation 4.25 changing its ellipticity. For our purposes we are free to assume this object is circularly symmetric in the source plane since the intrinsic ellipticity is degenerate with the effect of shear. In other words, instead of modelling shape in the usual way ($\varepsilon_{obs} = \langle \gamma, \varepsilon_{int} \rangle$), I take $\varepsilon_{int} = 0$ which makes γ equal to ε_{obs} with some random error. The result of this is that the brightness distribution of a galaxy in the source plane can now be described as a function of one variable, the distance from its centre ($|\mathbf{x}_s|$).

The shift in position of a galaxy due to weak lensing is small and unmeasurable; as such I ignore this effect and objects are assigned a position in the image plane after the shear field has been applied. This is done using a position vector coordinate transform. Hence, an unlensed galaxy at position $\mathbf{p}$ with profile f will have a brightness distribution in the image plane of,

$$I(\mathbf{x}) = f(|\mathbf{x}_s - \mathbf{p}|) \tag{5.1}$$

and an observation containing n such galaxies would be,

$$I(\mathbf{x}) = \sum_{g=1}^{n} f_g(|\mathbf{x}_s - \mathbf{p}_g|). \tag{5.2}$$

Adding shear

So far all galaxies in this model are unlensed i.e. $\mathbf{x} \equiv \mathbf{x}_s$. In order to include the lensing effect I define the functions $\gamma_1^T(\mathbf{x})$ and $\gamma_2^T(\mathbf{x})$ to be the true shear field across the image plane. These are combined with random noise $\varepsilon_{1,2}$ to imitate intrinsic ellipticity. The final transformation field in complex notation ($\gamma = \gamma_1 + i\gamma_2$) is then,

$$\gamma(\mathbf{x}) = \frac{\gamma^T(\mathbf{x}) + \varepsilon}{1 + \gamma^{T*}(\mathbf{x})\varepsilon} \tag{5.3}$$

In the weak lensing regime where $|\boldsymbol{\gamma}| < 0.1$ the shear over an individual galaxy can be assumed to be constant. As such I set the value of shear for a particular object (γ_g) to be that of the lensing field at the object's centroid and ε_g to be drawn at random such that $|\varepsilon|$ is normally distributed. Hence,

$$\gamma_g = \gamma(\mathbf{p}_g) = \frac{\gamma^T(\mathbf{p}_g) + \varepsilon_g}{1 + \gamma^{T*}(\mathbf{p}_g)\varepsilon_g} \tag{5.4}$$

This means that every object in the model is lensed and therefore it initially contains no foreground sources. However, due to its additive form, an unlensed galaxy can be added to the model by superimposing another with $\gamma^T = \mathbf{0}$. The source plane objects can now be lensed using a coordinate transform of the standard lensing Jacobian, $\mathbf{x} = \mathscr{A}\mathbf{x}_s$ where,

$$\mathscr{A}(\gamma) = \begin{pmatrix} 1-\gamma_1 & -\gamma_2 \\ -\gamma_2 & 1+\gamma_1 \end{pmatrix} \tag{5.5}$$

Thus, the total sky brightness becomes,

$$I(\mathbf{x}, \gamma) = \sum_{g=1}^{n} f_g(|\mathscr{A}_g\mathbf{x}_s - \mathbf{p}_g|) \tag{5.6}$$

Considering the case of a single source ($n = 1$) at the image centre ($\mathbf{p}_g = \mathbf{0}$). A Taylor expansion of I around $\gamma = 0$ shows that this can be decomposed into circular and non-circular components.

$$I(\mathbf{x}_{im}, \boldsymbol{\gamma}) \simeq I(\mathbf{x}, \mathbf{0}) + I_{\gamma_1}(\mathbf{x}, \mathbf{0})\gamma_1 + I_{\gamma_2}(\mathbf{x}, \mathbf{0})\gamma_2 + O(\gamma^2) \tag{5.7}$$

where I_γ denotes partial differentiation by γ.

$$I(\mathbf{x}_{im}, \boldsymbol{\gamma}) = f(|\mathbf{x}|) + \gamma_1 \left[\frac{\partial f}{\partial \gamma_1}\right]_{\boldsymbol{\gamma}=0} + \gamma_2 \left[\frac{\partial f}{\partial \gamma_2}\right]_{\boldsymbol{\gamma}=0} \tag{5.8}$$

$$I(\mathbf{x}_{im}, \boldsymbol{\gamma}) = f(|\mathbf{x}|) + \gamma_1 \left[\frac{\partial f}{\partial |\mathscr{A}\mathbf{x}|}\frac{\partial |\mathscr{A}\mathbf{x}|}{\partial \gamma_1}\right]_{\boldsymbol{\gamma}=0} + \gamma_2 \left[\frac{\partial f}{\partial |\mathscr{A}\mathbf{x}|}\frac{\partial |\mathscr{A}\mathbf{x}|}{\partial \gamma_2}\right]_{\boldsymbol{\gamma}=0} \tag{5.9}$$

This gives, (see appendix D for details)

$$I(\mathbf{x}_{im}, \boldsymbol{\gamma}) = f(|\mathbf{x}|) + \gamma_1 f'(|\mathbf{x}|)\frac{y^2 - x^2}{|\mathbf{x}|} + \gamma_2 f'(|\mathbf{x}|)\frac{-2xy}{|\mathbf{x}|} \tag{5.10}$$

$$I(\mathbf{x}_{im}, \boldsymbol{\gamma}) = f(|\mathbf{x}|) + f'(|\mathbf{x}|)\left(\frac{(y^2 - x^2)\gamma_1 - 2xy\gamma_2}{|\mathbf{x}|}\right) \tag{5.11}$$

This equation is now a function of $|\mathbf{x}|$ which is simpler if expressed in polar coordinates. Using the following transforms,

$$x = r\cos\theta \tag{5.12}$$

$$y = r\sin\theta \tag{5.13}$$

$$\Rightarrow |\mathbf{x}| = r \tag{5.14}$$

we obtain,

$$\Rightarrow I(r,\theta) \simeq f(r) - f'(r)\left(\frac{(\gamma_1 r^2(\cos^2\theta - \sin^2\theta)) + 2r^2\gamma_2\cos\theta\sin\theta}{r}\right) \tag{5.15}$$

$$I(r,\theta) = f(r) - f'(r)r(\gamma_1\cos 2\theta + \gamma_2\sin 2\theta) \tag{5.16}$$

Notice that the error on equation 5.7 and thus 5.16 is of order γ^2, which could be neglected if γ was dependent on weak lensing only. The inclusion of ε_{int} into this term means that this is not a valid assumption. However, the contribution from internal ellipticity is considered a source of noise which should be statistically removed. As such, it is not precisely clear how an estimator built from this approximation will be effected. It is likely that the estimator will be biased or have higher noise properties due to shape dispersion. Thus, there may be improvements which could be made by using a different model at this point.

Equation 5.16 describes a single lensed source in the image plane. It would be trivial to generalise this to n sources at position p; as has been shown previously using a coordinate shift and summation. Since the Fourier equivalent of these operations are well known however, I will first model the observation with this simple case and then generalise in Fourier space.

Radio Observation Transform

There are several parts to an interferometric radio observation that can affect the raw data; these are discussed in more detail in chapter 2. Firstly the response of individual antennas defines the reception pattern (B), next the correlation between elements produces data representing a Fourier transform of the received sky; this data is sparsely sampled by the function (S) defined by the array geometry. Finally there is observational noise such as antenna temperature, ionospheric disturbance and RFI (N). Given the sky brightness I, these effects produce the visibility set,

$$V(u,v) = N(u,v)S(u,v)\int B\frac{I(l,m)}{\sqrt{1-l^2-m^2}}e^{-2\pi i(ul+vm)}dldm \tag{5.17}$$

Usually these raw visibilities would be gridded, weighted and averaged to reduce data volume and improve computational efficiency. I will not consider these operations now as I am working with an analytical form of the data, but I later discuss the gridding process and the impact it may have on this method. I also absorb B and the w-projection terms into the sky model and disregard the telescope noise properties for this section since they should be stochastic and not systematically alter any shear signal. This gives a visibility model of a single source at the phase centre of,

$$V(u,v) \simeq S(u,v)\mathscr{F}_{lm}\{I(r,\theta)\} \tag{5.18}$$

where $\mathscr{F}_{lm}$ represents a Fourier transform with respect to the telescope coordinates l and m. Substituting in I, the model of sky brightness distribution from equation (5.16), this becomes,

$$S^{-1}(u,v)V(u,v) \simeq \mathscr{F}_{lm}\{f(r) - f'(r)r(\gamma_1\cos 2\theta + \gamma_2\sin 2\theta)\} \tag{5.19}$$

$$S^{-1}V = \mathscr{F}_{lm}\{f(r)\} - \mathscr{F}_{lm}\{f'(r)r(\gamma_1\cos 2\theta + \gamma_2\sin 2\theta)\} \tag{5.20}$$

Since the sky model is in polar coordinates we can express the visibilities in the Fourier equivalent, ρ and ϕ. The proper transforms in polar coordinates are given by Baddour (2009) resulting in,

$$S^{-1}V \simeq 2\pi\mathbb{H}_0\{f(r);2\pi\rho\} - \int f'(r)r^2(\gamma_1\cos 2\theta + \gamma_2\sin 2\theta)e^{-2\pi ir\rho\cos\theta-\phi}d\theta dr \tag{5.21}$$

Where $\mathbb{H}_0$ is a Hankel transform of order 0 defined in general as,

$$\mathbb{H}_\nu\{f(r);k\} = \int_0^\infty f(r)\,\mathscr{J}_\nu(kr)rdr \tag{5.22}$$

and $\mathscr{J}_\nu$ is a Bessel function which is commonly represented in integral form as,

$$\mathscr{J}_\nu(x) = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{i(\nu\tau - x\sin\tau)}d\tau \tag{5.23}$$

Since the function f is only dependent on r we can separate this from the integral over θ,

$$S^{-1}V \simeq 2\pi\mathbb{H}_0 - \int f'(r)r^2\int_{-\pi}^{\pi}(\gamma_1\cos 2\theta + \gamma_2\sin 2\theta)e^{-2\pi i r\rho\cos\theta-\phi}drd\theta \tag{5.24}$$

Evaluating the remaining expression over θ then results in a function of ϕ and a second order Bessel integral.

$$\frac{V}{2\pi S} \simeq \mathbb{H}_0 + (\gamma_1\cos 2\phi + \gamma_2\sin 2\phi)\int f'(r)r^2\,\mathscr{J}_2(2\pi\rho r)dr \tag{5.25}$$

Notice that the only part of the model now dependent on γ is the ϕ term preceding this integral which provides the main source of lensing distortion in the radio visibilities. I will now refer to this term as the lensing "kernel" defined by, $\mathbb{K}(\phi) = (\gamma_1\cos(2\phi) + \gamma_2\sin(2\phi))$. Concentrating on the remaining integral and applying a by parts method yields,

$$\int f'(r)r^2\,\mathscr{J}_2(2\pi\rho r)dr = \left[f(r)r^2\,\mathscr{J}_2(2\pi\rho r)\right]_0^\infty - \int f(r)\frac{d}{dr}\left[r^2\,\mathscr{J}_2(2\pi\rho r)\right]dr \tag{5.26}$$

$$\int f'r^2\,\mathscr{J}_2dr = -\int f(r)\frac{d}{dr}\left[r^2\,\mathscr{J}_2(2\pi\rho r)\right]dr \tag{5.27}$$

By applying a standard recurrence relation for Bessel functions, as given by Watson (1995),

$$\frac{d}{dr}\left[r^{\pm\nu}\,\mathscr{J}_\nu(kr)\right] = \pm kr^{\pm\nu}\,\mathscr{J}_{\nu\mp1}(kr) \tag{5.28}$$

(with $\nu = 2$ and $k = 2\pi\rho$) the derivatives in the expression can be removed leaving,

$$\int f'(r)r^2\,\mathscr{J}_2(2\pi\rho r)dr = -2\pi\rho\int f(r)r^2\,\mathscr{J}_1(2\pi\rho r)dr \tag{5.29}$$

$$\int f'r^2\,\mathscr{J}_2 dr = \rho\int f(r)r\left(-2\pi r\rho^{-0}\,\mathscr{J}_1(2\pi r\rho)\right)dr \tag{5.30}$$

Using the recurrence relation again but this time in the opposite direction and with $\nu = 0$, $k = 2\pi r$ and $r = \rho$ gives,

$$\int f'r^2\,\mathscr{J}_2 dr = \rho\int f(r)r\frac{d}{d\rho}\left[\rho^{-0}\,\mathscr{J}_0(2\pi r\rho)\right]dr \tag{5.31}$$

$$\int f'r^2\,\mathscr{J}_2 dr = \rho\frac{d}{d\rho}\int f(r)\,\mathscr{J}_0(2\pi r\rho)rdr \tag{5.32}$$

This is just a Hankel transform of order 0 from equation 5.22, and thus the polar Fourier transform of $f(r)$, thus,

$$\int f'(r)r^2\,\mathscr{J}_2(2\pi\rho r)dr = \rho\frac{d}{d\rho}\mathbb{H}_0 \tag{5.33}$$

$$\int f'(r)r^2\,\mathscr{J}_2(2\pi\rho r)dr = \rho\frac{d}{d\rho}\mathscr{F}\{f(r)\} \tag{5.34}$$

Substituting this back into the current visibility model in equation 5.25 gives the concise function,

$$\frac{V}{2\pi S} = \mathbb{H}_0 + \rho\mathbb{K}\frac{d\mathbb{H}_0}{d\rho} \tag{5.35}$$

or,

$$\frac{V}{S} = \mathscr{F}\{f\} + \rho\mathbb{K}\frac{d}{d\rho}\left[\mathscr{F}\{f\}\right] \tag{5.36}$$

So, from this we can see that the distortion due to galaxy ellipticity in the visibility domain can be simply expressed as the function $\rho\mathbb{K}(\gamma,\phi)$, multiplied by the radial derivative of the circular Fourier profile. This make sense intuitively since in physical space it translates to a convolution between the circular profile and lensing kernel.

Now this expression must be generalised to represent n objects at arbitrary positions. As in the earlier section a position can be represented as a coordinated change which in the Fourier is well known to be a phase shift. So, a galaxy with position (r_g,θ_g) can be represented in our model by,

$$V_g(\rho,\phi) = V_{centred} e^{-2\pi i \rho r_g \cos(\phi-\theta_g)} \tag{5.37}$$

An exception to this is the beam response which was absorbed into the galaxy profile. By assigning position in this way the beam shape is also re-centred on that galaxy's position. However, if the change in response pattern over the size of a galaxy is small, then it can be considered to be constant for that object. In this case we simply need to recall later that a flux of a galaxy's model should be weighted by the beam response in that direction.

Adding multiple galaxies to the model is done, as before, using a sum over the individual intensity. Since integration, and hence the Fourier transform, is distributive over addition this sum can be applied over the model in Fourier space. Thus our final model for an arbitrary number of lensed galaxies in the FOV of a radio telescope becomes,

$$\frac{V}{S} = \sum_{g=1}^{n} \left(\mathscr{F}\{f_g\} + \rho \mathbb{K}_g \frac{d}{d\rho} [\mathscr{F}\{f_g\}] \right) e^{-2\pi i \rho r_g \cos(\phi-\theta_g)} \tag{5.38}$$

Testing against numerical simulation

In order to characterize the error in our model due to the approximations made, I now compare this analytical expression to its numerical counterpart. In order to construct this I chose the galaxy profiles to be all Gaussian; this is not very realistic but is sufficient for the current purposes of studying the model. I generated 1000 galaxies over an area of $10x10$ arc-minutes giving a high source number density of $10 arcmin^{-2}$. For each object 6 random numbers are required to satisfy the following parameters,

- Position ($p_g = (p_{xg}, p_{yg})$): Given in arc-seconds as two uniformly distributed random numbers in the range $(-600, 600)$,
- Size (σ_g): Also provided in arc-seconds and drawn from the range $[1,5]$
- Peak flux (m_g): Since sensitivity is unimportant for this test, flux is a unit-less value drawn from a uniform distribution in the range $[1,10]$
- Shape ($\gamma_g = (\gamma_{1g}, \gamma_{2g})$): Since this test is not attempting to identify shear these parameters relate to ellipticity and are drawn from a normal distribution with $\bar{\gamma} = 0$ and $\sigma_\gamma = 0.3$.

Returning to the original model of sky brightness distribution this gives an expression which can be numerically evaluated,

$$I(x) = \sum_{g=1}^{n} m_g e^{|\mathscr{A}_g \mathbf{x} - p_g|^2/\sigma_g} \tag{5.39}$$

$$I(x) = \sum_{g=1}^{n} m_g e^{\left| \begin{array}{l} x(1-\gamma_{1g}) - \gamma_{2g} y - p_{xg} \\ y(1+\gamma_{1g}) - \gamma_{2g} x - p_{yg} \end{array} \right|^2 / \sigma_g} \tag{5.40}$$

$$I(x) = \sum_{g=1}^{n} m_g e^{\left((x-\gamma_{1g}x-\gamma_{2g}y-p_{xg})^2 + (y+\gamma_{1g}y-\gamma_{2g}x-p_{yg})^2\right)/\sigma_g} \tag{5.41}$$

The numerical visibilities V_{num} are then calculated using a Fast Fourier Transform (FFT) and the observation formula,

$$V = 2\pi \mathscr{F}\{B \times I\} \tag{5.42}$$

For this case a Gaussian is used for the beam response $B = e^{-(x^2+y^2)/150}$ so that toward the edge of the field $B \simeq 0.5$. The sampling function S is ignored since I am not applying a gridding function and the effect would be the same for V and V_{num}.

Substituting the Gaussian galaxy profile into the visibility model in equation 5.38 and adjusting the received flux using $B_g = B(p_g)$ gives,

$$V = \sum_{g=1}^{n} \left(\mathscr{F}\{B_g m_g e^{-r^2/\sigma_g}\} + \rho \mathbb{K}_g \frac{d}{d\rho} \left[\mathscr{F}\{B_g m_g e^{-r^2/\sigma_g}\} \right] \right) e^{-2\pi i \rho r_g \cos(\phi - \theta_g)} \tag{5.43}$$

Where the polar coordinates are related to Cartesian in the standard manner,

$$r = \sqrt{x^2 + y^2} \tag{5.44}$$

$$\tan\theta = \frac{y}{x} \tag{5.45}$$

$$\rho = \sqrt{u^2 + v^2} \tag{5.46}$$

$$\tan\phi = \frac{v}{u} \tag{5.47}$$

$$\tag{5.48}$$

Since B_g and m_g are constants for a particular g they can be removed from the Fourier transforms,

$$V = \sum_{g=1}^{n} B_g m_g \left(\mathscr{F}\{e^{-r^2/\sigma_g}\} + \rho \mathbb{K}_g \frac{d}{d\rho} \left[\mathscr{F}\{e^{-r^2/\sigma_g}\} \right] \right) e^{-2\pi i \rho r_g \cos(\phi - \theta_g)} \tag{5.49}$$

The remaining transform is now of a simple Gaussian which is well known to be another Gaussian scaled with σ,

$$V = \sum_{g=1}^{n} B_g m_g \left(2\pi\sigma_g e^{-\pi^2\sigma_g\rho^2} + \rho \mathbb{K}_g \frac{d}{d\rho} \left[2\pi\sigma_g e^{-\pi^2\sigma_g\rho^2} \right] \right) e^{-2\pi i \rho r_g \cos(\phi - \theta_g)} \tag{5.50}$$

$$V = \sum_{g=1}^{n} 2\pi\sigma_g B_g m_g \left(e^{-\pi^2\sigma_g\rho^2} + \rho \mathbb{K}_g \frac{d}{d\rho} \left[e^{-\pi^2\sigma_g\rho^2} \right] \right) e^{-2\pi i \rho r_g \cos(\phi - \theta_g)} \tag{5.51}$$

$$V = \sum_{g=1}^{n} 2\pi\sigma_g B_g m_g e^{-\pi^2\sigma_g\rho^2} e^{-2\pi i \rho r_g \cos(\phi - \theta_g)} \left(1 - \pi^2 \sigma_g \rho^2 \mathbb{K}_g \right) \tag{5.52}$$

This function can now also be evaluated numerically for discrete values of x, y, u and v. I used a grid of 3.6×10^7 points providing a resultant pixel size of 0.1" in observation space, and hence a maximum Fourier scale of $5 arcsec^{-1}$. Taking the absolute difference between the numerical and analytical models, I find $\sim$99 percent of pixels have a fractional error of less than 10^{-2} of any pixel within $5''$. This error typically peaks at $\sim$1 percent of a nearby source's flux. This is expected given a $O(\gamma^2)$ approximation and is unlikely to cause a serious difference in the lensing signal.

5.1.2 Constructing the Estimators

In this section my aim is to study the visibility model V and attempt to analytically extract quantities relating to γ. This extraction will then be used to define an operation $\langle V \rangle$ whose product will be directly proportional to either $\hat{\gamma}^T$ or $\bar{\gamma}^T$. These are the Fourier amplitudes of the true shear signal and the average over some region respectively.

In the following section I will use some contractions of the key elements in the visibility model to keep the discussion concise and simple. First, the Fourier transform of $f(r)$, the circular component of an object's profile will be defined as $E(\rho)$. Second, I will use $P(\rho, \phi)$ to refer to the additional phase from a source's position p_g so that, $P_g = e^{-2\pi i \rho r_g \cos(\phi - \theta_g)}$. Throughout this section I will continue to make contractions of expressions to keep them manageable. As such, I provide a list of these at the end of this section in table 5.1 for easy

reference. These contractions reduce the current form of the model in equation 5.52 to,

$$V = S\sum_{g=1}^{n} E_g P_g + \rho \mathbb{K}_g P_g \frac{dE_g}{d\rho} \tag{5.53}$$

I start by considering the case of Fourier amplitudes estimation. The function γ can be expressed in terms of its Fourier components ($\hat{\gamma}(\mathbf{l})$) through the transform relationship,

$$\gamma(\mathbf{x}) = \int_{-\infty}^{\infty} \hat{\gamma}(\mathbf{l}) e^{-2\pi i \mathbf{x}\cdot\mathbf{l}} dl \tag{5.54}$$

For an individual galaxy g this becomes,

$$\gamma(p_g) = \int_{-\infty}^{\infty} \hat{\gamma}(\mathbf{l}) e^{-2\pi i p_g\cdot\mathbf{l}} dl \tag{5.55}$$

Notice that the inversion of this expression would no longer be a integral but a sum over g since p_g is not a continuous quantity. Hence this relation is not exact but represents the best measurement of $\hat{\gamma}$ that can be made with a given galaxy distribution. Substituting this approximation into a single $\mathbb{K}_g P_g$ term in 5.53 and restricting ourselves to only the γ_1 component, i.e $\gamma_2 = 0$ we are left with,

$$\mathbb{K}_g P_g = \int \hat{\gamma}_1(\mathbf{l}) e^{-2\pi i p_g\cdot\mathbf{l}} dl \cos 2\phi e^{-2\pi i r_g \rho\cos(\phi-\theta_g)} \tag{5.56}$$

Converting P to Cartesian coordinates for a time,

$$\mathbb{K}_g P_g = \cos 2\phi \int \hat{\gamma}_1(\mathbf{l}) e^{-2\pi i p_g\cdot\mathbf{l}} e^{-2\pi i p_g\cdot\mathbf{u}} dl \tag{5.57}$$

$$\mathbb{K}_g P_g = \cos 2\phi \int \hat{\gamma}_1(\mathbf{l}) e^{-2\pi i (p_g\cdot\mathbf{l} + p_g\cdot\mathbf{u})} dl \tag{5.58}$$

Multiplying this by the complex conjugate of the position term (P_g^*) and the Fourier phase function $e^{2\pi i p_g\cdot\mathbf{l}_t}$ which I will now refer to as γ_t, then we find,

$$\mathbb{K}_g P_g^* \gamma_{tg} = \cos 2\phi \int \hat{\gamma}_1(\mathbf{l}) e^{-2\pi i (p_g\cdot\mathbf{l} + p_g\cdot\mathbf{u})} dl e^{2\pi i (p_g\cdot\mathbf{l}_t + p_g\cdot\mathbf{u})} \tag{5.59}$$

$$\mathbb{K}_g P_g^* \gamma_{tg} = \cos 2\phi \int \hat{\gamma}_1(\mathbf{l}) e^{-2\pi i (p_g\cdot\mathbf{l} + p_g\cdot\mathbf{u} - p_g\cdot\mathbf{l}_t - p_g\cdot\mathbf{u})} dl \tag{5.60}$$

$$\mathbb{K}_g P_g^* \gamma_{tg} = \cos 2\phi \int \hat{\gamma}_1(\mathbf{l}) e^{-2\pi i p_g\cdot(\mathbf{l}-\mathbf{l}_t)} dl \tag{5.61}$$

The term $e^{-2\pi i p_g \cdot (\mathbf{l}-\mathbf{l}_t)}$ is a delta function so the integral will give the value of $\hat{\gamma}$ at the "test" scale l_t.

$$\mathbb{K}_g P_g^* \gamma_{tg} = \cos 2\phi \hat{\gamma}(l_t) \tag{5.62}$$

However, this is only for the case where there is a single galaxy; returning to the full sum of n galaxies these individual terms cannot be separated out. Multiplying instead by the sum of $P_g^* \gamma_{tg}$ yields a slightly different answer,

$$\sum_{g=1}^{n} \mathbb{K}_g P_g \sum_{j=1}^{n} P_j^* \gamma_{tj} = \cos 2\phi \sum_{g,j}^{n} \int \hat{\gamma}(\mathbf{l}) e^{-2\pi i (p_g \cdot \mathbf{l} - p_j \cdot \mathbf{l}_t)} dl \tag{5.63}$$

$$\sum_{g=1}^{n} \mathbb{K}_g P_g \sum_{j=1}^{n} P_j^* \gamma_{tj} = \cos 2\phi \sum_{g,j}^{n} \int \hat{\gamma}(\mathbf{l}) e^{-2\pi i p_g \cdot (\mathbf{l}-\mathbf{l}_t)} e^{-2\pi i \Delta p_{gj} \mathbf{l}_t} dl \tag{5.64}$$

Again this is a delta function and hence $\hat{\gamma}(l_t)$ is recovered when $\Delta p_{gj} = 0$. For $g \neq j$ however we see more complex behaviour. For galaxy pairs that are separated by a distance approximating a multiple of the scale length l_t then $e^{-2\pi i \Delta p_{gj} \cdot \mathbf{l}_t} \simeq 1$ and it has little effect. For other pairs this term is less than 1 and damps the signal since these galaxies are not sensitive to that scale. This operation therefore produces a result that is proportional to $\hat{\gamma}(l_t)$ for each galaxy pair but weighted by $d_{gj}(l_t) = e^{-2\pi i \Delta p_{gj} \mathbf{l}_t}$.

$$\sum_{g=1}^{n} \mathbb{K}_g P_g \sum_{j=1}^{n} P_j^* \gamma_{tj} \propto \sum_{g,j}^{n} d_{gj} \cos 2\phi \hat{\gamma}(\mathbf{l}_t) \tag{5.65}$$

Returning to the full case model in equation 5.53 and applying this process to the second term results in,

$$\sum_{g,j}^{n} \rho \mathbb{K}_g P_g E_g' P_j^* \gamma_{tj} = \rho \left(\cos 2\phi \hat{\gamma}_1 + \sin 2\phi \hat{\gamma}_2\right) \sum_{g,j}^{n} d_{gj} E_g' \tag{5.66}$$

Which is a simple multiplicative relationship between $\hat{\gamma}$, galaxy profiles and the components of the kernel that I will now refer to as $\mathbb{K}_1 = \cos 2\phi$ and $\mathbb{K}_2 = \sin 2\phi$. If the quantity $\sum E_g P_g$ can be calculated or measured then the left hand side of this equation can be written in terms of V as,

$$\left(V - S \sum_g E_g P_g\right) \sum_j P_j^* \gamma_{tj} = \left(\hat{\gamma}_1 \mathbb{K}_1 + \hat{\gamma}_2 \mathbb{K}_2\right) S \rho \sum_{g,j} d_{gj} E_g' \tag{5.67}$$

Now the different γ terms need to be separated from each other and the dependence of the function on other variables such as ρ and ϕ removed. One way to collapse this function so that it is only dependent on l_t would be to integrate over ρ and ϕ. However, this would cause both the kernel components to average to zero and leave no shear signal to detect. Rather, if it is first multiplied by one of the lensing kernels that term will be positive and the integral will be a multiple of $\hat{\gamma}$.

$$\int \left(V - S\sum_g E_g P_g\right) \sum_j P_j^* \gamma_{tj} \begin{pmatrix} \mathbb{K}_1 \\ \mathbb{K}_2 \end{pmatrix} d\rho d\phi = \int \begin{pmatrix} \hat{\gamma}_1 \mathbb{K}_1^2 + \hat{\gamma}_2 \mathbb{K}_{1\times 2} \\ \hat{\gamma}_2 \mathbb{K}_2^2 + \hat{\gamma}_1 \mathbb{K}_{1\times 2} \end{pmatrix} S\rho \sum_{g,j} d_{gj} E_g' d\rho d\phi \tag{5.68}$$

Considering just the first of these (the multiple by $\mathbb{K}_1$), writing $\rho \sum_{g,j} d_{gj} E_g'$ as $F(\rho)$ since it is some arbitrary function of ρ and assuming a perfect sampling function ($S = 1$) gives,

$$\int \left(\hat{\gamma}_1 \mathbb{K}_1^2 + \hat{\gamma}_2 \mathbb{K}_{1\times 2}\right) F(\rho) dA = \hat{\gamma}_1 \int \mathbb{K}_1^2 d\phi F(\rho) d\rho + \hat{\gamma}_2 \int \mathbb{K}_{1\times 2} d\phi F(\rho) d\rho \tag{5.69}$$

$$\int \left(\hat{\gamma}_1 \mathbb{K}_1^2 + \hat{\gamma}_2 \mathbb{K}_{1\times 2}\right) F(\rho) dA = \left(\hat{\gamma}_1 \int_{-\pi}^{\pi} \cos^2 2\phi d\phi + \hat{\gamma}_2 \int_{-\pi}^{\pi} \frac{1}{2} \sin 4\phi d\phi\right) \int F(\rho) d\rho \tag{5.70}$$

$$\int \left(\hat{\gamma}_1 \mathbb{K}_1^2 + \hat{\gamma}_2 \mathbb{K}_{1\times 2}\right) F(\rho) dA = \pi \hat{\gamma}_1 \int F(\rho) d\rho \propto \hat{\gamma}_1 \tag{5.71}$$

Unfortunately, for an interferometer we cannot assume such a simple form for the sampling function, but we do know its exact form from the baseline positions. Therefore we can devise a weighting scheme that will give $\int_{-\pi}^{\pi} w(\rho,\phi) S \sin(4\phi) d\phi = 0$ as required by matching the distribution of $S\sin(4\phi)$ with $\sin(4\phi)$. This gives,

$$\int \left(\hat{\gamma}_1 \mathbb{K}_1^2 + \hat{\gamma}_2 \mathbb{K}_{1\times 2}\right) F(\rho) dA = \hat{\gamma}_1 \int \int_{-\pi}^{\pi} wS \cos^2 2\phi d\phi F(\rho) d\rho \tag{5.72}$$

$$\int \left(\hat{\gamma}_1 \mathbb{K}_1^2 + \hat{\gamma}_2 \mathbb{K}_{1\times 2}\right) F(\rho) dA = \hat{\gamma}_1 \int \int_{-\pi}^{\pi} \frac{1}{2} wS(\cos 4\phi + 1) d\phi F(\rho) d\rho \tag{5.73}$$

$$\int \left(\hat{\gamma}_1 \mathbb{K}_1^2 + \hat{\gamma}_2 \mathbb{K}_{1\times 2}\right) F(\rho) dA = \frac{1}{2} \hat{\gamma}_1 \int \int_{-\pi}^{\pi} wS \cos 4\phi + wS d\phi F(\rho) d\rho \tag{5.74}$$

Since cos is simply a shift of sin and $\int wS \sin 4\phi = 0$ then,

$$\int \left(\hat{\gamma}_1 \mathbb{K}_1^2 + \hat{\gamma}_2 \mathbb{K}_{1\times 2}\right) F(\rho) dA = \frac{1}{2} \hat{\gamma}_1 \int \int wSF d\phi d\rho \propto \hat{\gamma}_1 \tag{5.75}$$

The only part of the term multiplying $\hat{\gamma}$ that cannot be calculated exactly using the array configuration and galaxy positions is E'_g. It will be possible to estimate this from the data or calibrate the result. So, the final FILM estimator for Fourier estimation mode is given by $\langle V|l_t\rangle$ as the operation,

$$\langle V|l_t\rangle = \int\int \left(V - S\sum_g E_g P_g\right) \begin{pmatrix} \mathbb{K}_1 \\ \mathbb{K}_2 \end{pmatrix} \sum_j P_j^* \gamma_{tj} d\rho d\phi \tag{5.76}$$

$$\langle V|l_t\rangle = \begin{pmatrix} \hat{\gamma}_1(l_t) \\ \hat{\gamma}_2(l_t) \end{pmatrix} \frac{1}{2} \int\int w(\rho,\phi) S(\rho,\phi) d\phi \sum_{g,j} d_{gj} E'_g(\rho) \rho d\rho \tag{5.77}$$

In broad terms FILM, in Fourier estimation mode, creates a model of the data distortion due to a coherent lensing field on the scale l_t and sums the product of data and model. For observations which have a large shear signal on scales of l_t the model and data will agree, effectively squaring the data and producing a high output. Where model and data are less similar the product will be more random hence estimating a lower signal.

This method does require the source positions of galaxies in order to calculate the P and d_{gj} terms; statistics of the brightness distribution and size of objects could improve accuracy. As such, the method is not entirely separate of an imaging or cataloguing process. However, the information required is much less sensitive than shape to imaging and deconvolution methods, and these parameters would not require high fidelity to extract.

5.1.3 Pixel reconstruction mode

The second FILM mode that I investigate is more similar to a classic shear estimator in optical lensing. The aim of this pixel reconstruction mode is to produce an estimate of average shear for a pre-defined set of sub-regions (pixels) in the FOV.

Firstly, let the sources contained within the area of a single pixel be defined as a subset of all sources $j \in g$. Returning to the full estimator in equation 5.76 we know that,

$$\langle V|l_t\rangle_i = \int\int \left(V - S\sum_g E_g P_g\right) \mathbb{K}_i \sum_j P_j^* \gamma_{tj} d\rho d\phi \tag{5.78}$$

Consider the case of $l_t = 0$, this gives $\gamma_t = e^{2\pi i p_g \cdot 0} = 1$ and hence the estimator then becomes,

$$\langle V|0\rangle_i = \int\int \left(V - S\sum_g E_g P_g\right) w\mathbb{K}_i \sum_j P_j^* d\rho d\phi \tag{5.79}$$

$$\langle V|0\rangle_i = \int\int wS\rho \sum_{gj} E'_g P_g P_j^* \gamma_{ig} \mathbb{K}_i^2 d\rho d\phi \tag{5.80}$$

The PP^* term becomes the distance function d_{gj} from earlier but this time as a function of Fourier coordinates $d_{gj} = e^{-2\pi i \Delta p_{gj} \cdot \mathbf{u}}$ giving,

$$\langle V|0\rangle_i = \sum_{gj} \gamma_{ig} \int\int wS\rho d_{gj} \mathbb{E}'_g K_i^2 d\rho d\phi \tag{5.81}$$

For each galaxy pair the integral acts as a weight on γ_{ig}. Since the only terms that affect the value of the integral for different pairs of galaxies are E' and d_{gj} we can express this as,

$$\langle V|0\rangle_i = \sum_{gj} W(f_g(r), \Delta p_{gj}) \gamma_{ig} \tag{5.82}$$

We know that $d_{gj} = 1$ at $\rho = 0$. For nearby source pairs, where Δp_{gj} is small, d_{gj} will decrease slowly with ρ making the weight W smaller overall. As $\Delta p_g j$ increases, d_{gj} will oscillate rapidly with ρ causing the integral to average to zero for pair separations of more than a few galaxy radii. Hence, the weight W will be zero unless $0 \leq \Delta p_{gj} \lesssim 5''$ in which almost all cases $g = j$ giving,

$$\langle V|0\rangle_i = \sum_j W(f_j(r)) \gamma_{ij} \tag{5.83}$$

Thus, dividing $\langle V|0\rangle$ by the number of galaxies in subset j, and using a suitable weighting scheme, we obtain the average shear signal in the pixel that defined j.

5.2 Simulations

In this section we provide details of the simulations created in order to test the FILM estimator. Our aim was to create a set of semi-realistic observations including the majority of problems a real radio lensing survey would face. We also wished to keep the simulations simple enough to fully understand, easy and fast to manage, and be of comparable quality to

Parameter	Definition	Description
$E_g(f(x))$	$2\pi\mathbb{H}\{f_g(x), 2\pi x'\}$	Fourier transform of galaxy profile f_g
$P_g(p_g, x)$	$e^{-2\pi i p_g \cdot x}$	Fourier phase from galaxy position p_g
$\mathbb{K}_g(x)$	$\gamma_{1g}\cos 2x + \gamma_{2g}\sin 2x$	Kernel of lensing for galaxy g
$\mathbb{K}_i(x)$ i=1,2	$(\cos 2x, \sin 2x)$	Separate kernels of shear for γ_i
$\gamma_t(x)$	$e^{2\pi i p_g \cdot x}$	Fourier mode testing function
$d_{gj}(x)$	$e^{2\pi i \Delta p_g \cdot x}$	Fourier phase separation function
$F(x)$	$x\sum_{gj} d_{gj} E'(x)$	Temporary reference for simplicity
$w(x)$	$\int w(x)S(x)\sin 4x dx = 0$	Weighting function to reduce sampling effect.

Table 5.1 Reference table of the various expression contractions used in this chapter for convenience.

next-generation instruments capable of detecting a weak lensing signal.

5.2.1 True Sky Model

We begin with the analytical expression, from which FILM is derived, of a field of sources in equation 5.6. This is a simple sum over n sources centred at position $\mathbf{p}_g$ with a shear assigned according to each source's centroid.

The galaxy population is drawn from a random patch of the s-cubed catalogue detailed in Wilman et al. (2008). This catalogue defines 4 types of radio component, cores, hot-spots, lobes and disks which are combined in various ways to create different radio structures. For these simulations we select all components (unless stated otherwise), within some distance of a randomly selected point (sky_{cen}), down to a flux limit of $1\mu Jy$. This limit gives an average source density of $\sim 10\text{arcmin}^{-1}$. For each component we create a model as described below and add it to a sky model at its position offset (p_g) from sky_{cen}.

Cores and Hot-spots:
Components labelled as cores and hot-spots can be modelled as point sources. As such the 1.4GHz total flux is added to the pixel closest to $\mathbf{p}_g$. Since we are not considering the change in flux or number density due to lensing, no shear is included in their model.

Disks:
At the low flux limits ($\sim \mu Jy$) needed to reach a high number density of sources, star-forming and starburst galaxies form an increasing proportion of the radio population (Simpson et al. (2006); Wilman et al. (2008)). We model these objects using an exponential profile (sersic

with n=1) with amplitude (m) drawn from the 1.4GHz peak flux, and size (σ) given by the average of the major and minor axes. This is then given shape and shear as described in section 5.1.1 using the lensing Jacobian ($\mathscr{A}$).

$$I_g = m_g \cdot e^{-\mathscr{A}_g(|\mathbf{x}|)/\sqrt{\sigma_g}} \tag{5.84}$$

$$I_g = m_g \cdot e^{-((x-\gamma_{1g}x-\gamma_{2g}y)^2+(y+\gamma_{1g}y-\gamma_{2g}x)^2/\sigma_g)^{\frac{1}{2}}} \tag{5.85}$$

This is then shifted to the correct position and added to the sky model using the co-ordinate transform,

$$\mathbf{x} \rightarrow \mathbf{x}_g = \mathbf{x} - \mathbf{p}_g \tag{5.86}$$

For simulations that include intrinsic ellipticity ε this is added to the shear as described in section 5.1.1, using a Gaussian distribution with $\sigma_\varepsilon = 0.3$.

Lobes:
A sersic profile can also be used as a simple model for the lobe components which make up galaxies with a FR type morphology. As such we add these components in the same way as disks, with the exception of intrinsic ellipticity. Since lobes of the same galaxy are radiated from a central core they tend to be aligned with each other and have a similar ellipticity (Saripalli, 2012). For a pair of lobe components we assign the same intrinsic ellipticity using a position angle that orients each lobe towards the centroid of the other. The magnitude of ε is still drawn from the same $\sigma_\varepsilon = 0.3$ Gaussian distribution used for disks. This model is a little extreme as most FR type galaxies will not be this intrinsically aligned, but it does serve as an interesting test case.

For some tests included in the results section we omit most component types including only disk galaxies; we do consider the impact of these other sources in section 5.3.3.

5.2.2 Lensing signal

We choose a "thin-sheet" approximation to the shear field, i.e. the lensing potential is considered to be at some fixed point from the observer ($z = d_l$). Hence, all objects with $z < d_l$ are unlensed and thus referred to as "foreground sources". Galaxies at a distance greater than $z = d_l$ are all lensed by $\gamma(x)$ at their position regardless of redshift. This approach does disregard the change to lensing potential over redshift, and also neglects its correlation to

foregrounds and large scale structure. Most reconstructions shown in the results section are not from a realisation that includes foreground sources with the exception of those in section 5.3.3. In general d_l was chosen to be at $z = 0.2$ since a typical radio survey will include galaxies up-to $z \sim 1$.

FILM is designed to reconstruct Fourier modes of lensing shear. As such, we generate the γ maps used in equation 5.85 from the Fourier domain. This was done in one of two ways; single-mode and multi-mode. Single-mode shear fields are created by Fourier inverting a single pixel representing a delta function. This results in an oscillating shear pattern on some chosen scale. This case is not very realistic but is useful for illustration. Figure 5.5 shows an example of a γ_1 map in observation space generated using the single-mode method. Multi-mode shear maps represent a more realistic signal. Amplitudes for scales representing one arc-minute or larger are generated from a normal distribution with Hermitian symmetry. Scales less than an arc-minute are set to zero. In physical space this corresponds to a randomly fluctuating real valued γ map with features on arc-minute scales and larger. This is then normalised such that in real space $\sigma_\gamma = 0.1\frac{\ln 2}{2} \simeq 0.35$, i.e $\gamma < 0.1$ for 99.5% of the sample.

5.2.3 Simulated observation

The gridded visibilities for an observation of the above simulations are created using a Fourier Transform. The sky model itself is generated on a grid, and we are therefore able to employ an FFT to minimize the complexity of this step. In order to replicate the effects that the gridding would have on real visibilities, we simulate the sky model to higher resolution and larger fov than needed for the specified observation. The exact choices for the simulation and motivation are described below.

Beam Response

The antenna response pattern or beam $B(r)$ is modelled as a Gaussian weighted sinc function of the offset from pointing centre. We include scale factors such that the FWHM of the primary beam defines the field of view (FOV). This beam model is a simplified version of the response given by a dish with an unblocked aperture, for example, the offset Gregorian used by meerKAT (Booth et al., 2009), and eventually SKA-mid. The beam model $B(r)$ is

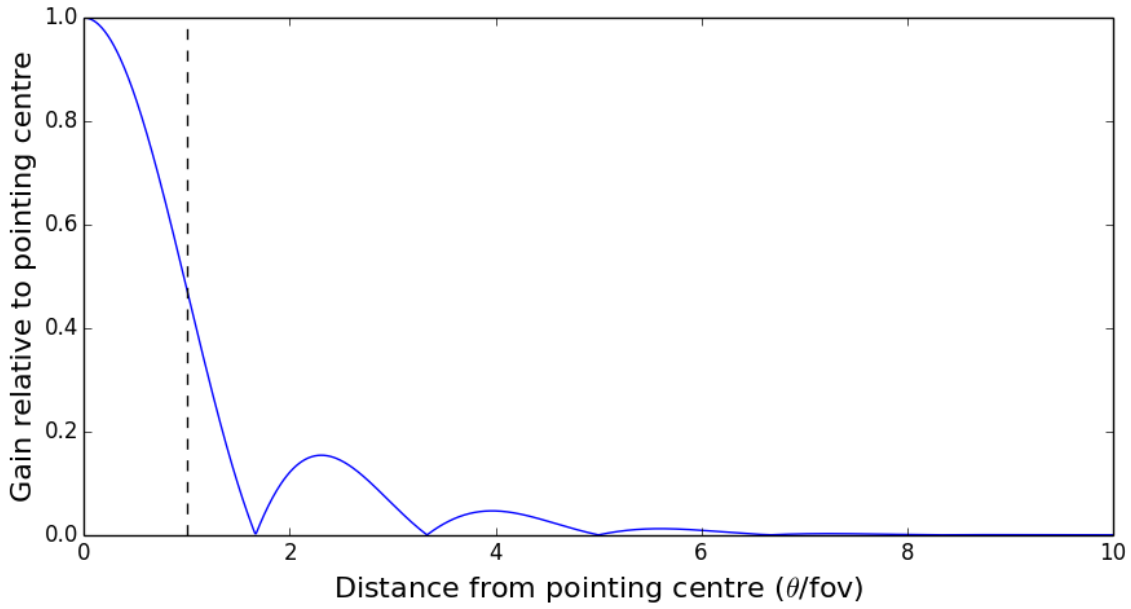


Fig. 5.2 Beam model for simulated instrument. Black dashed line shows the fov at primary beam FWHM.

defined as,

$$B(r) = e^{-\left(\frac{r}{4FOV}\right)^2} \cdot \left| sinc\left(\frac{0.6r}{FOV}\right)\right| \tag{5.87}$$

A plot of this model can be seen in figure 5.2. Notice that the response beyond the second side lobe is less than $\sim 1\%$ of that at the pointing centre and thus has a negligible effect. As such we ignore this region by only simulating the sky model to 5 times the required FOV. Multiplying the sky model by B results in the received intensity,

$$I = B(|\mathbf{x}|)\sum_g m_g \cdot e^{-\mathscr{A}_g(|\mathbf{x}_g|)/\sqrt{\sigma_g}} \tag{5.88}$$

A Fourier transform of I calculated using an FFT will give a set of gridded visibilities oversampled by a factor of 5,

$$V(u,v) = FFT\{I(x,y)\} \tag{5.89}$$

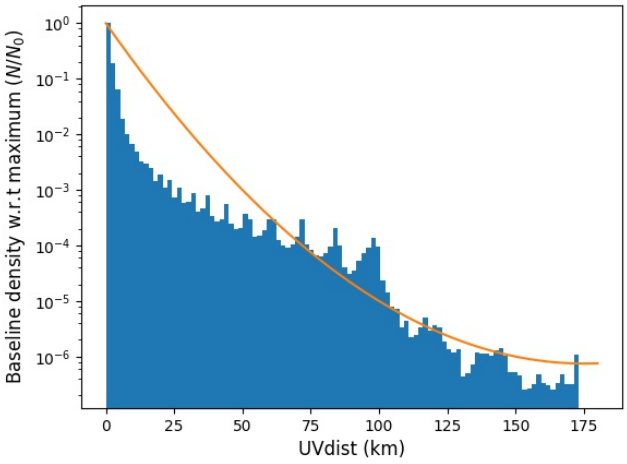


Fig. 5.3 UV coverage density. Blue histogram is simulated SKA-mid baseline count as a function of UVdistance normalised by area and the centre bin count. Orange line shows logarithmic baseline weighting used for the simulation.

UV Coverage and Antenna Noise

For most cases the signal is reduced in the Fourier domain to replicate the sparse coverage of arrays using SKA1 mid at 1.4GHz as a model. The longest baselines of the SKA1-mid are expected to be 150km (Garrett et al., 2010); at 1.4GHz this is equivalent to $\sim 0.3''$ resolution. I simulate the initial sky model at $0.25''$ giving measurable visibilities up to 180km.

Reduction of the data quality due to UV coverage is simulated in one of three ways. Firstly, an annulus function (uvA) giving no baselines at $|u| < 0.02$km or $|u| > 50$km and full coverage elsewhere. This region corresponds to the area covered by 90 percent of baselines in a 30 minute SKA1-mid observation and reduces the overall resolution but does not produce an anisotropic PSF. Secondly, a weighting of noise on long baselines in line with the fall in baseline density for SKA-mid (uvR). The baseline density is modelled as $10^{2\times10^{-4}|u|^2-7\times10^{-2}|u|}$, with baselines at $|u| < 0.02$km or $|u| > 180$km again removed. A comparison of the fall in density for a 30 minute SKA1-mid track with this weighting is shown in Figure 5.3. Noise is then added to gridded visibilities scaled by $1/\sqrt{uvR}$. This affects small scales most hence reducing fidelity, but again does not produce an ellipticity distorting PSF.

Finally, I use a simulated SKA coverage, as seen in figure 5.4 (provided by P. Patel from the SKA group at Manchester). For simulations with no telescope noise this is simply

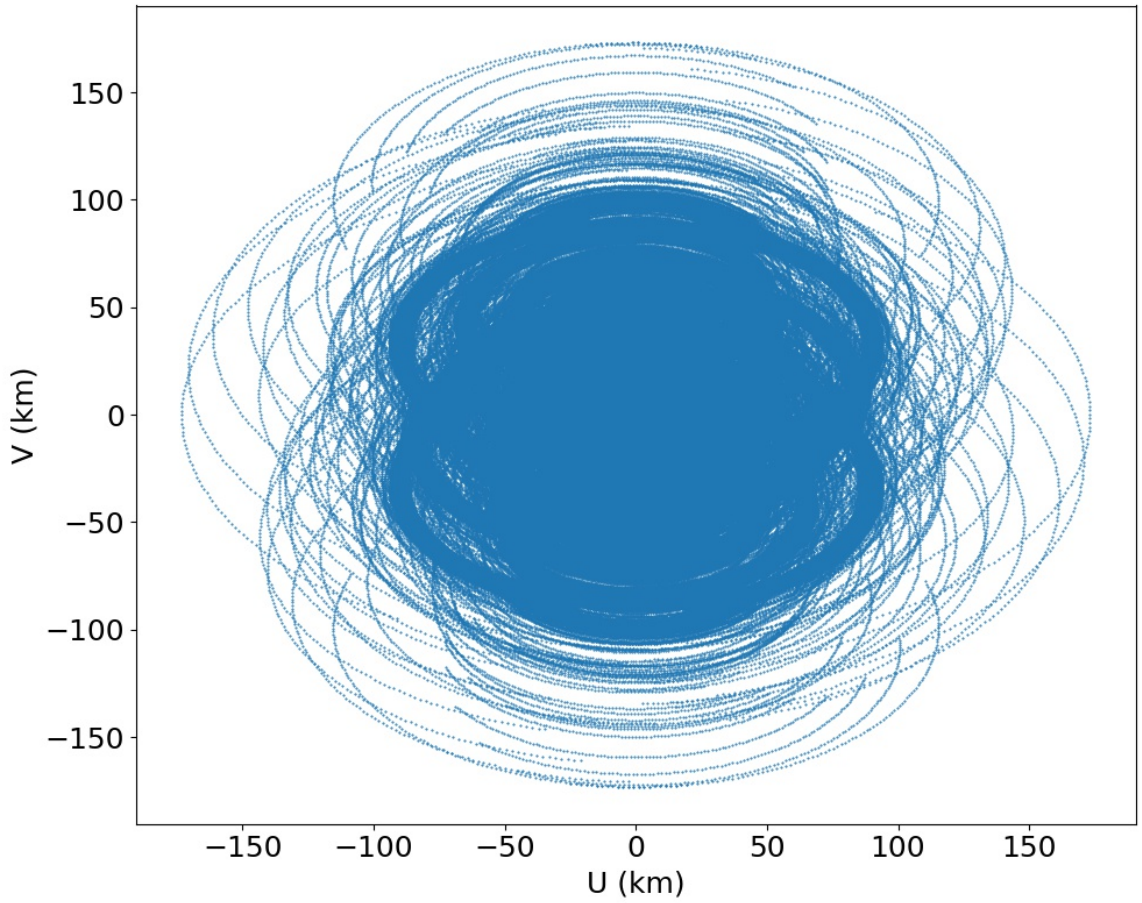


Fig. 5.4 Simulated UV coverage track for SKA1-mid pointing to zenith. 30 minute observation at 1.4GHz with integration over 15 seconds. Total of 7×10^6 interferometric pairs.

implemented as a weighting of one for pixels containing at least one baseline point and zero otherwise. When observational noise is included the noise is weighted by $\frac{1}{\sqrt{N}}$ for a pixel containing $N > 0$ UV points. This is similar to using a uniform weighting scheme as described in chapter 2 which gives the best resolution at the cost of point source sensitivity. Antenna noise, when included, is Gaussian and normalised such that the rms in real space is approximately 2×10^{-7}Jy, i.e the faintest included sources will have a peak flux 5σ above the noise.

Finally, the data must be regridded down from the $5\times$ oversampling that was used to incorporate the beam sidelobes. Since this is an exact multiple of the required gridding, I am able to use a fast convolution algorithm, rather than the slow direct convolution that would be needed for raw data. I use a sinc function as the gridding kernel, described in section 2.4.2,

convolve with the data and average over every 5 pixels.

Summary:

- *Sources*: All simulations contain lensed disk like galaxies with exponential profiles at a distance of $z > 0.2$. Some additional tests also include: point sources, such as hot-spots and cores; large-scale galaxies comprised of two lobes; and unlensed foreground sources at $z < 0.2$, which all contribute noise to the signal. The parameters of all these objects, position, size and peak flux at 1.4GHz, are extracted from four square degrees of the Wilman et al. (2008) model including components down to threshold 1μJy.

- *Source density*: The exclusion of sources with peak flux $< \mu$Jy or $z > 0.2$ provides an average source density of $\sim 10\text{arcmin}^{-1}$.

- *Lensing signal*: For most cases $\hat{\gamma}^T$ is generated as a randomly fluctuating field with a variance of $\sigma_{\gamma^T} \simeq 0.03$ on arc-minute scales. An exception to this is the first case shown which is defined by a delta function in Fourier space resulting in the sinusoidal pattern for γ^T seen in figure 5.5.

- *Ellipticity noise*: Normally distributed random quantity with variance $\sigma_\varepsilon = (0, 0.1, 0.3)$, added to γ_g^T on an individual galaxy basis using the expression in equation 5.3 giving γ_g.

- *Resolution*: Model provides measurable visibilities generated up to Fourier scales equivalent to $0.25''$ at 1.4GHz. This just exceeds the capabilities of SKA1-mid. The resolution is then affected by the chosen sampling function.

- *Beam response*: The *sinc* reception pattern shown in figure 5.2 is applied to all simulations.

- *Sampling*: Two sampling functions are used for the main results: an annulus function (uvA) which removes all baselines outside $|u| \in (0.02, 50)$km; and a parabolic, logarithmic weighting *uvR*, also zero for $|u| \notin (0.02, 180)$km. The effect of sampling function based on an SKA simulated track is also considered in section 5.3.3.

- *Instrumental Noise*: Antenna noise is generated from a normal distribution in Fourier space and normalised such that the rms of the difference of the sky map in observation space is $\simeq 2 \times 10^{-7} Jy$.

5.3 Results

In this section I show the results from applying FILM to simulations as described previously with a variety of lensing signals, noise levels and intrinsic shape distributions. For each case we show the correlation between the true shear field (γ_{in}) and the FILM estimation (γ_{out}), and the variance of the difference ($\sigma^2_{\delta\gamma}$). In addition to results based on the basic simulations I also discuss how decreased positional accuracy, a realistic UV-coverage and the inclusion of a non-disk-like foreground population affect the measurement. The results described in this section are summarised in table 5.2.

As seen in equation 5.75 the result of the FILM estimator is not directly the γ or $\hat{\gamma}$ field. Rather, FILM provides a result which is proportional to the desired measures of shear and a further nationalisation must be applied. The form of this normalisation can also be seen in equation 5.77. It is a complex function of the source characteristics (size, brightness and distribution), the total number of *uv* visibilities, and the residuals of the weighted sampling function (*wS*). It may be possible to characterize this normalisation using only the visibility data, otherwise a source extraction technique will be needed. Further work is required in exploring the form of this normalisation in order to produce reliable results. As such, I have used the best possible linear fit for the results shown in the following section. This means that although these results are a good representation of the noise from the current estimator, they are in no way indicative of the bias.

5.3.1 Fourier Estimation (FE) mode

Single Fourier mode shear field:

When operating in FE mode, FILM estimates the amplitude of specified Fourier scales. As such an interesting first test is to consider an unrealistic shear field described by a single Fourier scale, i.e, a delta function. The estimation that FILM provides near this scale will describe the "leakage" from strong amplitudes. As such, the first lensing signal reconstruction we show is of a δ function in Fourier space at

$$\ell = \begin{pmatrix} 0.005 \\ 0.005 \end{pmatrix} arcsec^{-1} \tag{5.90}$$

for $\widehat{\gamma_1}$, and

$$\ell = \begin{pmatrix} -0.005 \\ 0.005 \end{pmatrix} arcsec^{-1} \tag{5.91}$$

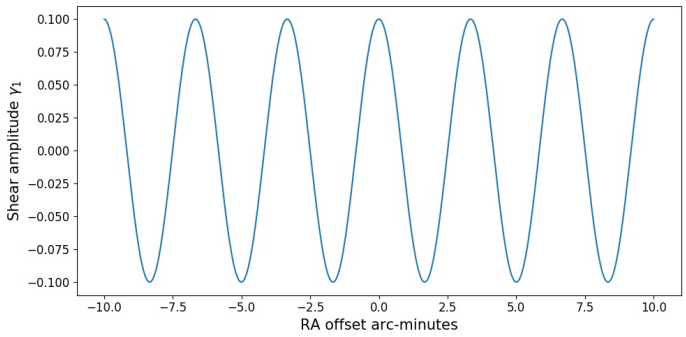


Fig. 5.5 Slice of the γ_1 signal map for the single Fourier mode case at a declination offset of zero.

for $\widehat{\gamma_2}$.

These amplitudes represent a sinusoidal fluctuation in the shear signal on scales of ~ 1.5 arc-minutes. In real space this shear field is a uniform diagonal wave, the amplitude map of which can be seen in figure 5.5 for γ_1. For this example I have used the uvA sampling function, temporarily increasing the maximum baseline cut-off to 150km. The simulation used includes no intrinsic ellipticity, no telescope noise and no foreground or point source objects.

FILM reconstruction was performed on scales up to $\ell = 10^{-2} arcsec^{-1}$ with a step-size of $\ell_{pix} = 10^{-3}$. This sampling is equivalent to measuring scales from 50arcsecs up to 8arcmins. Since this case contains no signal other than at $\ell = 0.005$ it is important that the reconstruction grid includes this point. However, this prior knowledge will not be necessary for typical reconstructions where $\hat{\gamma}$ is smoother. Figure 5.6 shows this reconstruction and the difference map with $\hat{\gamma}^T$; a clear peak can be seen at the true Fourier mode location with SNR≈ 5.

A Fourier transform of the result in figure 5.6 provides a low resolution estimate of the true shear field. For this case, where no telescope or shape noise was included in the simulation, the correlation between the mean of γ_{in} over the reconstructed pixels and γ_{out} has a Pearson value of $r^2 = 0.89$.The distribution of the estimated shear error $\delta\gamma = \gamma_{in} - \gamma_{out}$ can be seen in Figure 5.7 with a variance of $\sigma^2_{\delta\gamma} = 3.6 \times 10^{-4}$.

So far we have considered a simulation where the shape of galaxies is due only to cosmic shear and are circular otherwise. We now include a random intrinsic ellipticity field as described in section 5.2 with variance $\sigma_\varepsilon = 0.3$. As expected this increases the noise of the

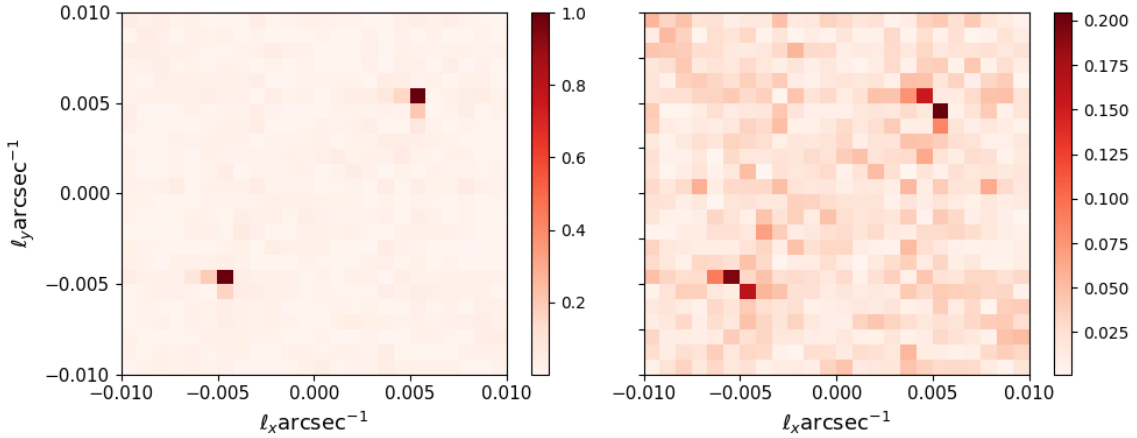


Fig. 5.6 Left: FILM estimate of single mode $|\hat{\gamma}_1|$ signal normalised by peak amplitude. Right: Residual of $|\hat{\gamma}_{1in} - \hat{\gamma}_{1out}|$ with the same normalisation.

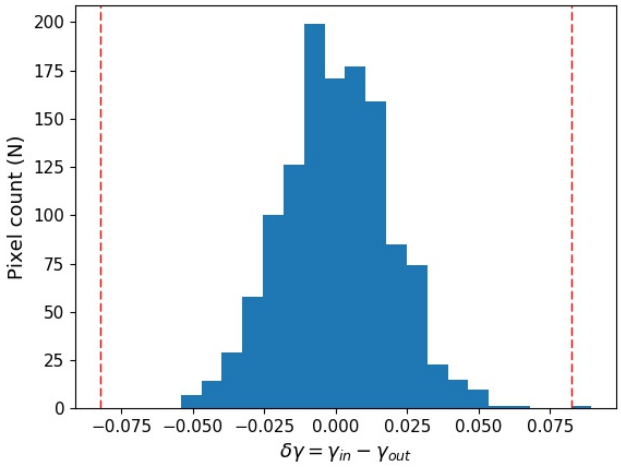


Fig. 5.7 Histogram shows the distribution of the difference between true and reconstructed shear ($\delta\gamma = \gamma_{in} - \gamma_{out}$), for a simulation with no noise. The variance of the distribution is $\sigma^2_{\delta\gamma} = 3.6 \times 10^{-4}$. Red lines show minimum and maximum values of γ_{in}.

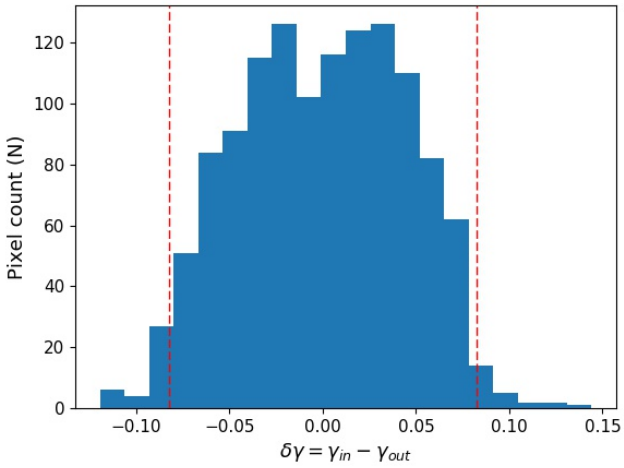


Fig. 5.8 Histogram shows distribution of difference between true and reconstructed shear ($\gamma_{in} - \gamma_{out}$), for a simulation with ellipticity noise of $\sigma_\varepsilon = 0.3$. The variance of the distribution is $\sigma^2_{\delta\gamma} = 2 \times 10^{-3}$. Red lines show minimum and maximum values of γ_{in}.

final reconstruction, with the correlation coefficient between γ_{in} and γ_{out} pixels reducing to $r^2 = 0.41$. Again the distribution of errors is shown in figure 5.8; here we can see that the error variance increases by a factor of five, with $\sigma^2_{\delta\gamma} = 2 \times 10^{-3}$.

This reconstruction was performed using a high resolution of 50" and we can reduce the Gaussian shape noise on small scales by smoothing. Using a top-hat filter we down-weight scales smaller than 1'. This improves the correlation coefficient to $r^2 = 0.58$ and the variance of the residuals to $\sigma^2_{\delta\gamma} = 1.4 \times 10^{-3}$.

The results above show that this method can identify a single large scale shear mode present in radio visibilities with a realistic background source population even on a small field. I now test FILM on shear fields which are more realistic.

Multi-mode shear field

The results in this section are based on different simulations increasing in complexity and noise and include the random Gaussian lensing signals described in section 5.2.2 with σ_γ =0, 0.1 and 0.3. For each case I show the correlation between the "true" shear field, and the best normalization to the FILM estimate. This correlation is shown for both Fourier

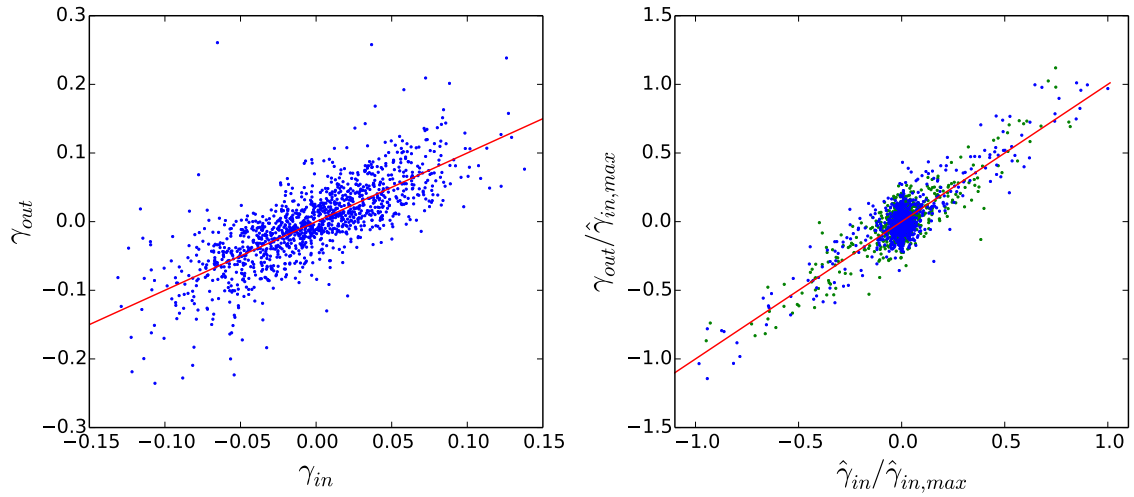


Fig. 5.9 Correlation of γ_{in} vs γ_{out} for a simulation with: $\sigma_\varepsilon = 0$, uvA sampling and no telescope noise. Right panel shows Fourier amplitudes ($\hat{\gamma}$) with the real part in blue and imaginary in green: Pearson value $r^2 = 0.72$, Variance $\sigma^2_{\delta\gamma} = 0.02$. Left panel shows real space shear (γ) from a Fourier inversion of the right. Pearson value $r^2 = 0.59$, Variance $\sigma^2_{\delta\gamma} = 0.001$. Red line is $\gamma_{in} = \gamma_{out}$.

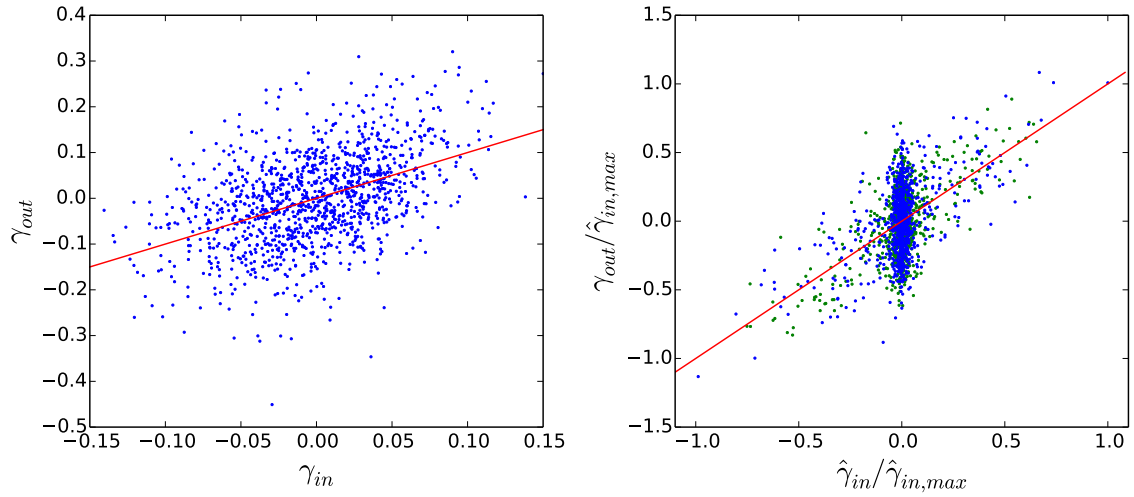


Fig. 5.10 Correlation of γ_{in} vs γ_{out} for a simulation with: $\sigma_\varepsilon = 0.3$, uvA sampling and no telescope noise. Right panel shows Fourier amplitudes ($\hat{\gamma}$) with the real part in blue and imaginary in green: Pearson value $r^2 = 0.31$, Variance $\sigma^2_{\delta\gamma} = 0.01$. Left panel shows real space shear (γ) from a Fourier inversion of the right. Pearson value $r^2 = 0.21$, Variance $\sigma^2_{\delta\gamma} = 0.008$. Red line is $\gamma_{in} = \gamma_{out}$.

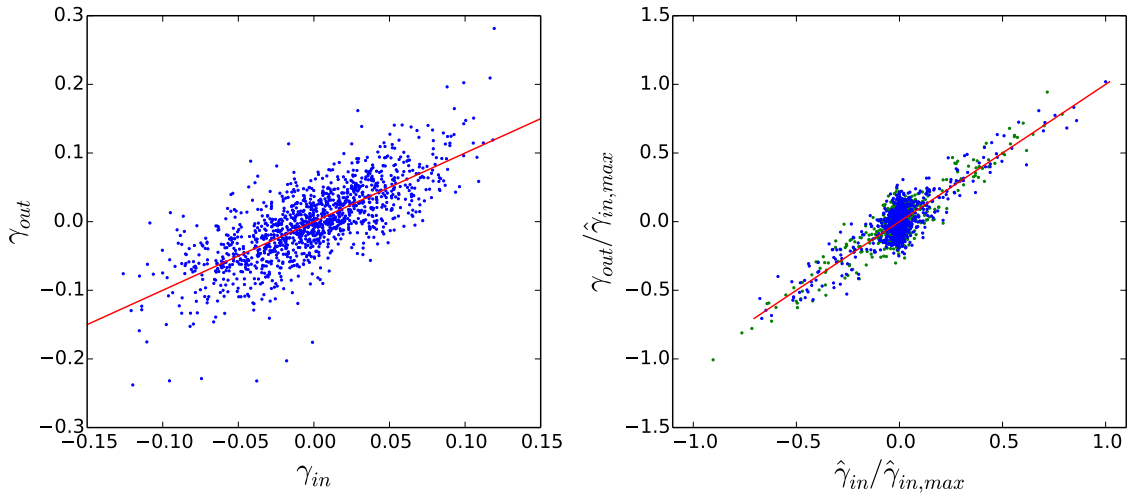


Fig. 5.11 Correlation of γ_{in} vs γ_{out} for a simulation with: $\sigma_\varepsilon = 0$, uvR sampling and telescope noise of rms=2×10^{-7}. Right panel shows Fourier amplitudes ($\hat{\gamma}$) with the real part in blue and imaginary in green: Pearson value $r^2 = 0.75$, Variance $\sigma^2_{\delta\gamma} = 0.015$. Left panel shows real space shear (γ) from a Fourier inversion of the right. Pearson value $r^2 = 0.56$, Variance $\sigma^2_{\delta\gamma} = 0.001$. Red line is $\gamma_{in} = \gamma_{out}$.

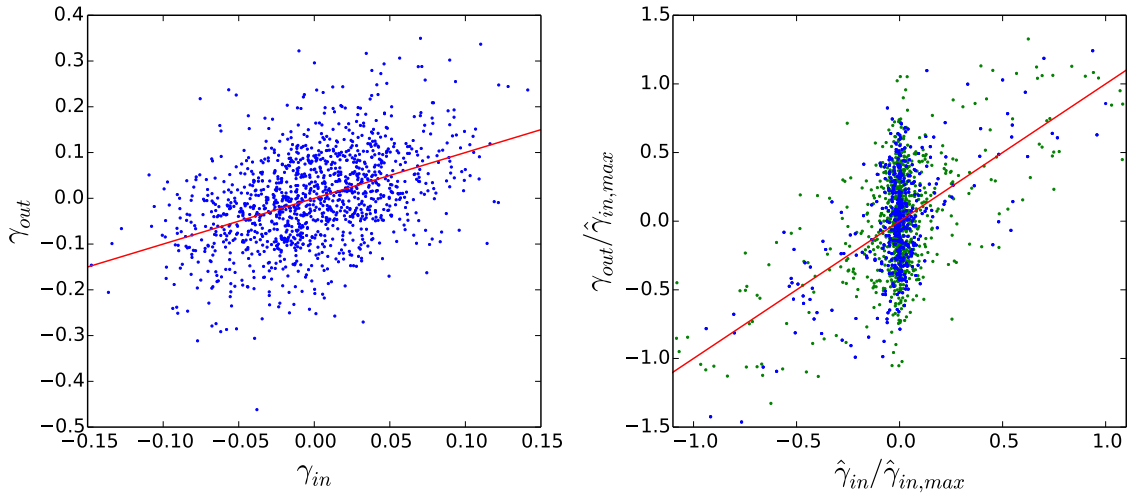


Fig. 5.12 Correlation of γ_{in} vs γ_{out} for a simulation with: $\sigma_\varepsilon = 0.3$, uvR sampling and telescope noise of rms=2×10^{-7}Jy. Right panel shows Fourier amplitudes ($\hat{\gamma}$) with the real part in blue and imaginary in green: Pearson value $r^2 = 0.32$, Variance $\sigma^2_{\delta\gamma} = 0.22$. Left panel shows real space shear (γ) from a Fourier inversion of the right. Pearson value $r^2 = 0.19$, Variance $\sigma^2_{\delta\gamma} = 0.009$. Red line is $\gamma_{in} = \gamma_{out}$.

amplitudes and their transformed real space counterparts. I state the Pearson r^2 value for each case and variance of the difference between the true and normalised FILM estimate, $\sigma^2_{\delta\gamma}$.

Starting with the most basic case, Figure 5.9 shows the result from a simulation including no intrinsic ellipticity or telescope noise and using the uvA sampling. There is an excellent agreement between the true and estimated values with a Pearson value of $r^2 = 0.72$ and residual variance of $\sigma^2_{\delta\gamma} = 0.02$ in Fourier space. Note that in Fourier space $\hat{\gamma}$ is normalised to the range [-1,1], so this is a small residual. The agreement of the inversion is less clear with a larger error variance. This may in part be due to the small scale (50") to which it is generated. By using a more aggressive smoothing this error could be reduced, although this data is limited by the small (20 arc-minute) size of the simulation.

Increasing the intrinsic ellipticity noise to the typical value of $\sigma_\varepsilon = 0.3$ gives the result in 5.10. Some correlation can be seen in Fourier space but this is overwhelmed by scatter due to shape noise. By increasing the smoothing scale to $100''$ however, we can improve the result finding $r^2 = 0.63$ and $\sigma^2_{\delta\gamma} = 0.097$.

I also performed an estimate using an intermediate ellipticity scatter of $\sigma_\varepsilon = 0.1$ this resulted in $r^2 = 0.61$ and $\delta\gamma = 0.055$ in Fourier space. Hence the error due to intrinsic ellipticities is simply increasing as expected with shape dispersion. It can therefore be reduced using the usual approaches of smoothing and careful weighting.

For the next cases I move onto simulations that include telescope noise and use the uvR sampling scheme. This produces higher noise for visibilities in regions with lower baseline density. There is still a clear correlation of $r^2 = 0.75$ in figure 5.11 based on a simulation with $\sigma_\varepsilon = 0$. This increases when $\sigma_\varepsilon = 0.3$ as shown in figure 5.12 but by no more than in the simulation with no instrument noise. Again I was able to reduce this by smoothing on $100''$ scales giving a result of $r^2 = 0.64$ and $\sigma^2_{\delta\gamma} = 0.21$. Hence this type of random instrumental noise on the level expected for SKA1-mid would not cause significant problems for this method.

5.3.2 Pixel Reconstruction (PR) mode

In this section we test the performance of the alternative FILM mode that reconstructs a shear map on an individual pixel basis. For this we use a Gaussian lensing signal as in the previous

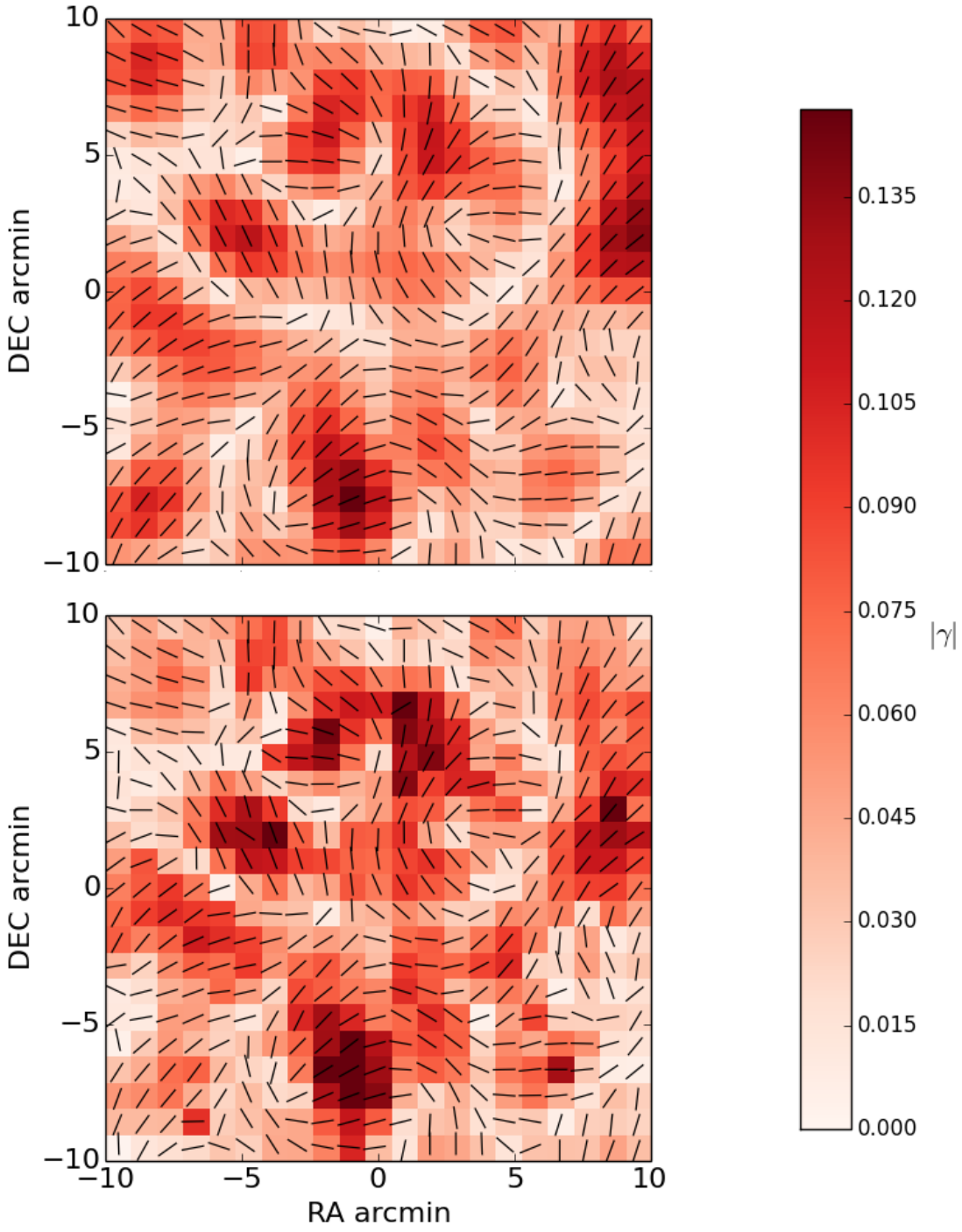


Fig. 5.13 True γ_{in} shear field top, FILM reconstruction from simulation with: $\sigma_\varepsilon = 0$, uvA sampling and no telescope noise, bottom. Vector lines show position angle PA$= \frac{1}{2}\arctan\left(\frac{\gamma_2}{\gamma_1}\right)$, colour map indicates $|\gamma| = \sqrt{\gamma_1^2 + \gamma_2^2}$.

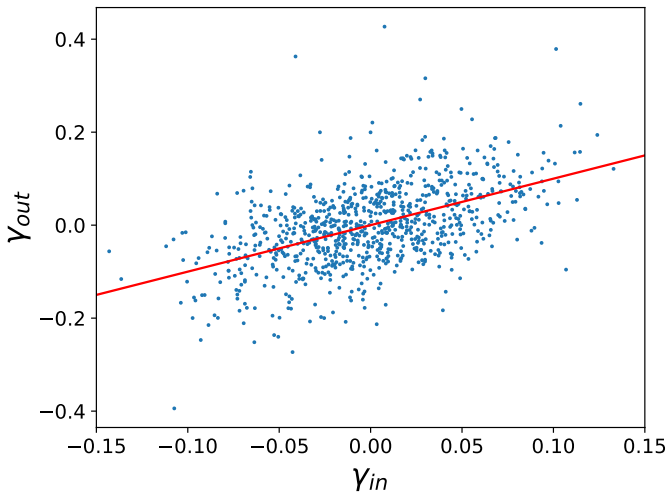


Fig. 5.14 Correlation of γ_{in} vs γ_{out} for FILM in PR mode using a simulation with: $\sigma_\varepsilon = 0.3$, uvR sampling and telescope noise of of rms=2×10^{-7}Jy. Pearson value $r^2 = 0.26$, Variance $\sigma^2_{\delta\gamma} = 0.006$. Red line is $\gamma_{in} = \gamma_{out}$.

section and show correlations for different simulations with increased noise.

Starting again with the most basic simulation containing no telescope or shape noise, a clear correlation can be seen between the true shear map and the FILM reconstruction in figure 5.13. These two maps have a correlation value of $r^2 = 0.76$. This is comparable to the value seen for this case using FE mode in Fourier space, but it performs better than a reconstruction of the inversion of FE mode estimate. This may be due to variations in number density which are accounted for in PR mode but not FE. Hence there is some improvement that can be made to FE mode by using the source distribution, for example, convolving with inverse galaxy density.

Returning to a more realistic case, figure 5.14 shows the result of applying PR mode FILM to simulations using, $\sigma_\varepsilon = 0.3$, uvR sampling and including telescope noise. This results in a correlation of $r^2 = 0.26$ with an error variance of $\sigma^2_{\delta\gamma} = 0.006$. Although this correlation is worse than the equivalent case in FE mode, it again performs better than the FE mode inversion. Again by using a more aggressive smoothing of $100''$, I improve this to $r^2 = 0.62$ and $\sigma^2_{\delta\gamma} = 1.7 \times 10^{-4}$.

5.3.3 Future considerations

In this section I discuss the next steps for FILM. I consider some of the further sources of noise which would be present in a real survey, and would potentially effect this estimator in particular. Additionally, I reflect the some of the gaps between FILM and current techniques, discussing the feasibility of developing FILM to into a competitive method.

Normalisation and bias correction

As discussed, most array configurations are not circularly symmetric. This isotropy would cause a severe distortion to the shear measurement if not corrected. In 5.1.2 I described a possible weighting scheme for the visibilities in order to reduce this effect in FILM. These weights were designed such that,

$$\int w(\rho,\phi)S(\rho,\phi)\sin 4\phi d\rho d\phi = 0 \tag{5.92}$$

and the distribution of $S\sin 4\phi$ was a close as possible to $\sin 4\phi$. This process seeks not to perform a traditional deconvolution of the PSF from the whole data set, but rather correct its effect on the shear signal only. This should result in an additional, but deterministic, multiplicative bias of the form,

$$m = \int wSE' d\rho d\phi. \tag{5.93}$$

This correction along with the discrete nature of the calculation means that a further normalisation of the result from FILM is required. Using the *uv* coverage seen in figure 5.4 as the sampling function I am still working to produce a weighting that gives a consistent and robust result. This is mainly due to the complex interplay between the sampling function and the gridding process.

The removal of bias is a process which will be vital for any future lensing surveys. Without the correct normalisation it will be impossible to characteristic and calibrate the bias of FILM. Furthermore, if the bias is still dependent on the brightness and position distribution of sources after this normalisation then calibration may still not be possible. Therefore, the investigation of the normalisation is the key next step in developing FILM.

Positional errors

Since FILM relies heavily on using well measured positions of source galaxies one may be concerned about the effect of astrometric error. Current positional errors in surveys such as

LOFAR can be as high as a few arc-seconds which would be enough to destabilise the d_{gj} term in equation 5.76 severely impacting any signal. However, there are several reasons why this should not be a problem in future.

Firstly, while current LOFAR images have a resolution of $\sim 10''$, this looks set to fall in the future to sub-arc-seconds with positional errors to match. Secondly, even if source positions in images are not reliable due to the non-linear imaging processes, radio provides other methods for measuring exquisite positions through fringe fitting, Cohen & Shaffer (1971). Finally, it is possible to use high precision optical positions for many compact radio objects. This has the added potential of using source positions which are below the sensitivity of the radio telescope. The shear signal from these sources would not be usable in a traditional individual ellipticity approach, but here any signal will be already combined in the Fourier domain contributing to the estimation.

Large Un-lensed foreground sources and multi-component morphologies

So far the simulations used have only contained lensed disk galaxies at a distance of $z > 0.2$. A real survey however would contain foreground sources which are typically brighter and larger than more distant objects. The visibility data would also include point sources, such as stars and hot-spots, the lobes of FR type galaxies which have a large intrinsic alignment signal, and other diffuse sources that are not a useful source of shear signal. All these additional objects will add noise to the estimator.

We are currently working to test the impact of foreground objects. The simple approach to this is to supply only the positions of the lensed background galaxies for use in the 'j' subset of the FILM operation. A more advanced method would be to model and attempt to remove these sources from the data first. Alternatively, we can restrict ourselves to the PR mode where we can explicitly mask out regions near bright sources.

3D lensing

Most future weak lensing studies will attempt to constrain cosmic shear in three dimensions i.e. including distance along the line of sight. This will allow weak lensing experiments to be more sensitive probes of dark energy, and the equation of state. There are two main methods to extending lensing measurements into a third dimension. One approach is to split the galaxy sample into tomographic "slices" based on their redshift (Hildebrandt et al.,

2017). Since there is a linear relationship between source distance and the distribution of the lensing signal, each slice will probe a different redshift range. The auto and cross-correlation between these slices can then be measured and used to constrain different models. A more advanced approach is to reconstruct the full 3D power spectra using spherical harmonics as presented by Kitching et al. (2014).

Uncertainties in redshift measurements, and the insufficient cosmological volumes currently surveyed, mean that 3D cosmic shear is not yet competitive with other probes (Kilbinger, 2015). However, the next generation of instrumentation will provide the required data quality. The inclusion of some redshift information will therefore be essential for any competitive technique.

A tomographic approach may also be possible with FILM. As mentioned in the previous section subsets of galaxies can be probed by the estimator by only providing the position information of those sources. Selecting several subsets of galaxies in different redshift slices should return the desired cosmic shear estimates dependent on distance. As with foreground removal this would need to be tested and is likely to require modifications to the estimator.

It may also be possible to construct a FILM-like estimator for full 3D estimation. However, this would require a complete recalculation of the initial model in three dimensions. It would also introduce at least two new parameters increasing the already high computational workload.

5.4 Conclusions

Radio weak lensing can provide the cosmology community with a new and complementary source of information if we are able to successfully extract the signal. In this chapter I have considered the form of the lensing shear signal in radio data and developed a novel method for modelling and measuring it.

I have demonstrated that given a high quality observation set, it is possible to estimate both a map of shear and its Fourier transform directly from the visibilities. I have also shown that for a wide survey the error on this measure can be reduced using larger smoothing scale. This is much the same for traditional optical lensing.

I have identified areas that may affect the performance of FILM on a real radio weak lensing survey and discussed the impact each of them. FILM is unlikely to suffer in future due to source position error, and in fact the inclusion of sub-noise sources may even improve the signal. Countering the effects of an anisotropic PSF and foreground sources is possible, but requires further investigation. The most important next step in developing FILM is to characterise the normalisation of the shear estimates. This will allow the noise and bias of FILM to be properly measured and compared to other methods.

FILM is a simple and efficient method which proves that shear analysis can, in principal, be performed on an entire radio observation. There are many ways in which improvements could be made to the FILM estimator, such as the inclusion of source morphology, redshifts or position distribution information. However, this work has formed a proof of concept for direct visibility lensing measurements.

5.4.1 Summary table of results

Simulation			Smoothing	Fourier mode result		Pixel mode result	
UVcov	N	σ_ε		$\sigma^2(\hat{\gamma}_{in}-\hat{\gamma}_{out})$	r^2	$\sigma^2(\gamma_{in}-\gamma_{out})$	r^2
uvA	×	0	$50''$	0.02	0.74	6.7×10^{-4}	0.76
			$100''$	0.04	0.86	2.8×10^{-5}	0.9
			$150''$	0.04	0.88	1.2×10^{-5}	0.92
		0.1	$50''$	0.05	0.61	0.001	0.63
			$100''$	0.08	0.81	2.6×10^{-5}	0.85
			$150''$	0.002	0.86	1.7×10^{-5}	0.82
		0.3	$50''$	0.1	0.31	0.006	0.26
			$100''$	0.097	0.63	1.3×10^{-4}	0.72
			$150''$	0.095	0.71	5.4×10^{-5}	0.73
uvR	✓	0	$50''$	0.015	0.75	0.001	0.59
			$100''$	0.018	0.9	2.6×10^{-5}	0.88
			$150''$	0.02	0.93	1.2×10^{-5}	0.86
		0.1	$50''$	0.036	0.63	0.001	0.59
			$100''$	0.04	0.84	3.8×10^{-5}	0.86
			$150''$	0.04	0.88	2.3×10^{-5}	0.84
		0.3	$50''$	0.22	0.32	0.006	0.26
			$100''$	0.21	0.64	1.8×10^{-4}	0.62
			$150''$	0.25	0.67	10^{-4}	0.56

Table 5.2 Summary of FILM results. UV coverage schemes described in 5.2.3. Noise defined such that $I_{rms} = 2\times10^{-7}$Jy. FE mode statistics calculated in Fourier space. For reference $\hat{\gamma}_{in}$ is normalized to $[-1,1]$, and γ_{in} is typically in the range $(0.1,0.1)$. All of these results use simulations of 320 square arc-minutes ($\sim$0.1 square degrees).

Chapter 6

Further Analysis Techniques and Conclusions

In this thesis we have seen that the sensitivity, resolution and speed of radio telescopes has improved exponentially over recent years. This trend looks to continue in the future with ever larger and more advanced instrumentation planned, most notably the SKA. The dramatic increase in data volume and quality has begun to outstrip the analytic techniques available to us. Without new efficient methods of analysis we risk losing valuable discoveries in the future. The principal aim of this thesis was to investigate the techniques that would benefit emerging radio astronomy/cosmology science aims.

I have so far presented two projects motivated by ambitious next generation survey plans. Firstly, a pilot study for the upcoming WEAVE-LOFAR programme (Smith et al., 2016) aiming to spectroscopically follow-up more than 10^6 low-frequency selected radio sources. This survey will produce a data set with a broad set of primary science cases whilst leaving enormous legacy value. My work has used current LOFAR data together with ancillary BOSS observations in order to maximise the efficiency of this future programme. Secondly, the development of a new radio weak lensing estimator. It has been shown previously that detection of a weak lensing signal at radio wavelengths is possible (Cheng, 2005; Patel et al., 2010). More recently Harrison et al. (2016) has demonstrated that next generation instruments will be competitive and highly complementary to optical and IR surveys. Indeed, if current projections of resolution and sensitivity are correct then perhaps both regimes will be considered equally important in the future. However, this is highly dependent on our ability to extract the lensing signal from the large sets of complex radio data.

6.1 Further Work

In addition to this work I have also undertaken several further projects that are relevant to this thesis. In this section I will give a brief overview of this additional work and discuss further research that is required in these areas to benefit radio cosmology. I will then provide an overview of the thesis with the conclusions of my work in the context of radio cosmology, finishing with a final summary.

6.1.1 Imaging of Wide Field LOFAR Data

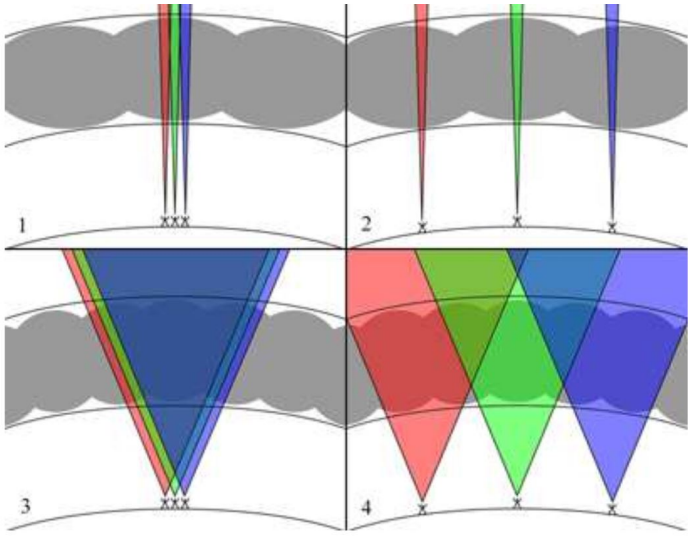


Fig. 6.1 The effect of ionosphere on various types of radio configuration: 1. Short baselines, narrow FOV, 2. Long baselines, narrow FOV, 3. Short baselines, wide FOV, 4. Long baselines, wide FOV. Reproduced from Intema et al. (2009).

As I have discussed in section 2.4.2, creating science quality images from radio data is a long complex process with multiple non-linear steps and subjective options. One element that can cause serious problems is the distorting effect of the ionosphere. This was mentioned previously in reference to optical lensing where the effect is removed through AO and point source calibration. At longer wavelengths atmospheric diffraction can be even more pronounced but is generally counteracted using a similar calibration technique. This works well for arrays with short baselines and a narrow FOV, as seen in panel 1 of figure 6.1, since the effect is uniform for all antennas and directions. Other array configurations, such as those with a large FOV or long maximum baselines, will require separate calibration for distant elements or directions. We call this direction dependent calibration.

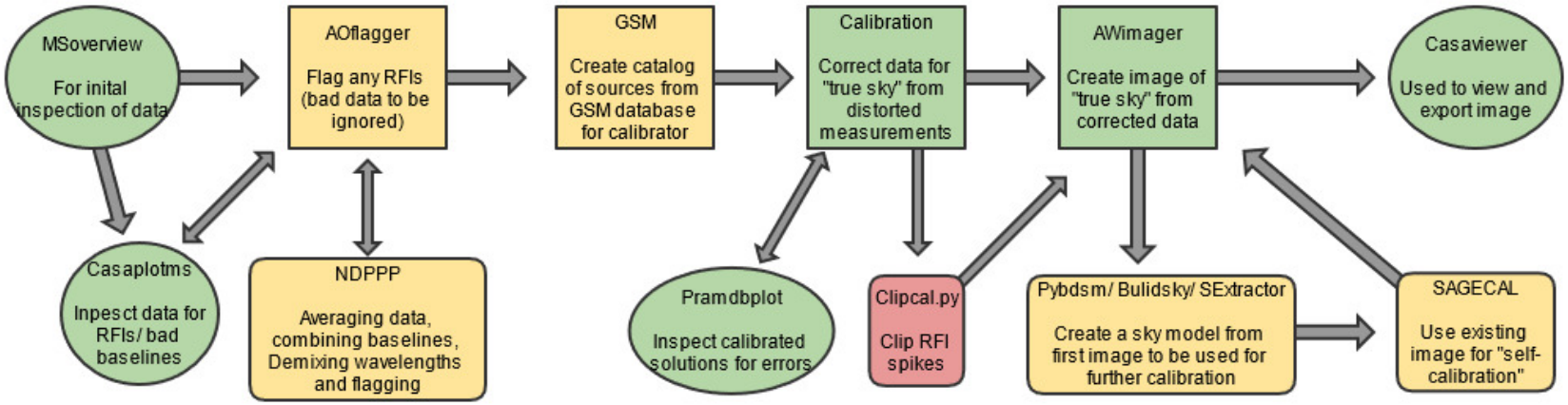


Fig. 6.2 Key stages and software programmes in the imaging pipeline when under construction. Main process steps shown in top row. Green shows components that were working and fully automated, yellow highlights steps where there remained bugs in the automatic process and red indicates a broken step.

Procedures such as direction dependent calibration or self-calibration are extremely time consuming and the way in which they are applied is often based on examining an initial low resolution image. Even this step can be resource intensive, especially for the very large observations currently being produced by LOFAR. As part of the LOFAR "Blank Fields" working group in the Surveys KSP I investigated implementing an automatic pipeline on SCIAMA at Portsmouth that would quickly create a simple first image to help this process and provide computational relief.

Figure 6.2 shows a schematic of the various scripts and programmes combined into the pipeline that would produce an image from a radio measurement set (MS). The main process steps are shown in the first row with optional steps below. The operations are as follows:

MSoverview: This is an application within the Common Astronomy Software Applications package (CASA), (McMullin et al., 2007)) that produces a report on the MS and checks for errors or corrupted data. The pipeline reads the output of MSoverview checking for a text string that confirms the data are "fully regular" before proceeding. If this string is not detected the output of MSoverview is returned instead.

AOflagger: A tool developed by Offringa et al. (2012) to remove RFI present in the data set based on unlikely peaks in the received amplitudes over frequency and time. After making the MS set writeable the pipeline implements AOflagger using default settings. Based on other studies this should remove $\sim$ 95percent of RFI from a typical observation.

CASAplotms (optional): This CASA application is typically used interactively to create plots of the data that can be used to identify bad baselines or RFI. If used this option will provide plots of amplitude vs time, amplitude vs frequency and visibility phase vs time along with the final image.

NDPPP (optional): The "New Default Preprocessing Pipeline" can perform many operations on the MS set to prepare it for imaging. In this pipeline the main process which is used is averaging. It is possible to continue without averaging the data but this makes the imaging process extremely long. If this option is used then NDPPP is run using a parset which averages the MS in frequency to one channel, but not in time.

Another key operation that can be performed using NDPPP is demixing. This involves subtracting the contribution of a strong nearby source that may bleed into the rest of the

image. As this pipeline is only producing an initial image this step is not used. At the time of creating this pipeline the AOflagger was incorporated into NDPPP. As such, these steps can now be combined which would be faster and less resource intensive compared to executing the steps separately.

Global Sky Model (GSM): Once the data has been averaged and flagged it is ready for phase calibration. In order to perform this calibration a model of known compact objects is required to compare the phases against. This is obtained by querying the GSM that contains objects from the VLSS, NVSS, and WENSS survey catalogues.

Calibration: The program BlackBoard Selfcal (BBS) is used for the initial phase calibration attempting to minimize the difference between the observed MS and the GSM model. BBS is then run on the data, first solving for the difference with the GSM model and then correcting the data. A direction independent solution is found here since dependent calibration is very lengthy.

Parmdbplot (optional): This script, if enabled, creates additional plots of the data in phase and amplitude vs time after calibration so that the solutions can be inspected. These should be smooth if the calibration was successful.

AWimager:At this point the phase solutions should have been successfully applied and an image can be made. AWimager is a bespoke imaging tool in CASA for LOFAR that performs both w-projection and A-projection. The pipeline runs the imaging process at 30" resolution with 6000 iterations and using a briggs weighting scheme with robust=0 (Briggs, 1995). This should result in a low resolution image that isn't over fitted to noise. Since no amplitude calibration was performed the fluxes will be incorrect, but it will provide quick insight into any issues such as bright sources near the edge of the FOV.

CASA viewer: Finally the CASAviewer program is used to created fits files of the image and deconvolution residuals.

Self calibration steps (advance): In order to produce a higher quality image the user may specify a number of iterations of self calibration. This is performed using the catalogue tool PyBDSM (Mohan & Rafferty, 2015) and the SAGECAL algorithm described in Kazemi et al. (2011). Using the initial low resolution image a model is made from the output of PyBDSM. SAGECAL then solves and corrects the data based on this model and a new higher

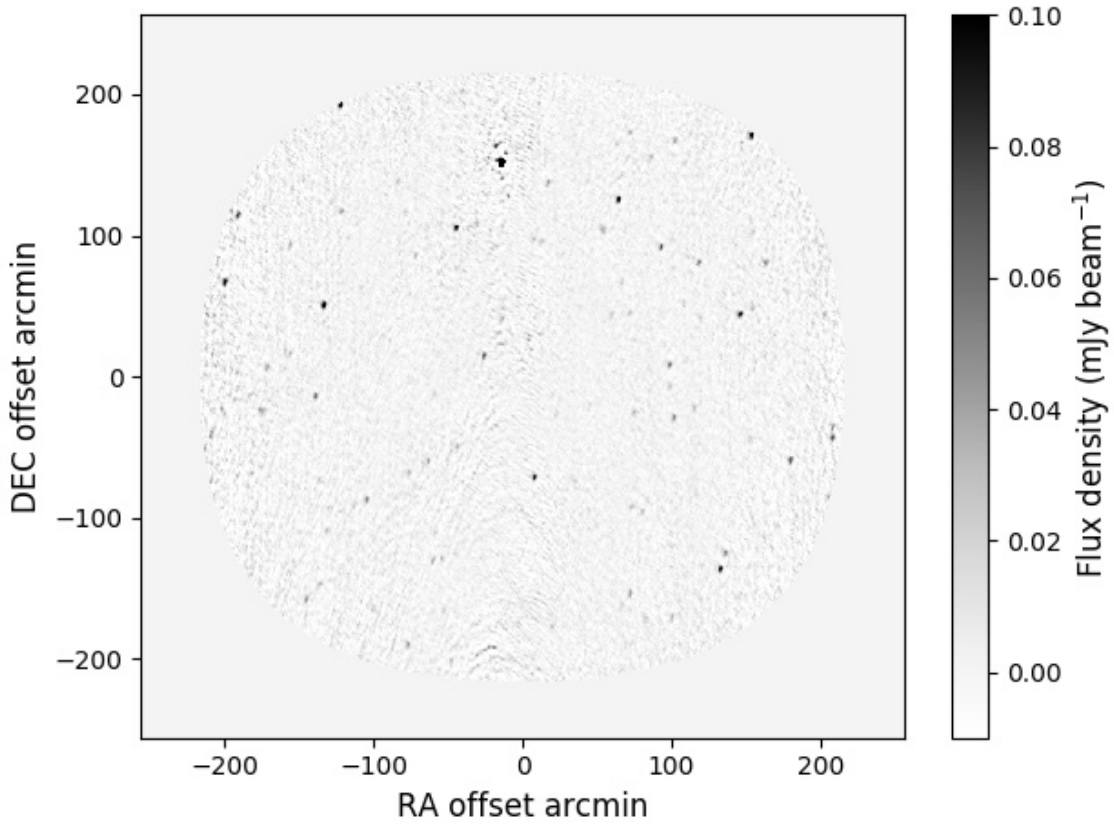


Fig. 6.3 Image of the XMM field from a single sub-band LOFAR observation produced using this pipeline. The fluxes in this image are unreliable since only phase calibration has been performed. The field covers 35 square degrees.

resolution image can be made. This process is prone to extrapolating error if not used with careful masking techniques and hence is notoriously difficult to automate. It is included in the pipeline as an experimental feature.

I worked to build this pipeline, deciding on the parameters and the calibration strategy used. I then tested the pipeline producing images of several different observations. Figure 6.3 shows an image of the XMM field produced using this pipeline. There is a large amount of noise ($rms = 0.07\text{mJybeam}^{-1}$) due to the bright objects, as seen in figure 6.4. Solutions for this include peeling and direction dependent calibration Pandey et al. (2009).

Recent observation such as those of the Boötes field have used a new calibration technique called facet calibration developed by Williams et al. (2016). In this method the FOV is broken into regions called facets that are defined using a Voronoi tessellation of calibrator sources in

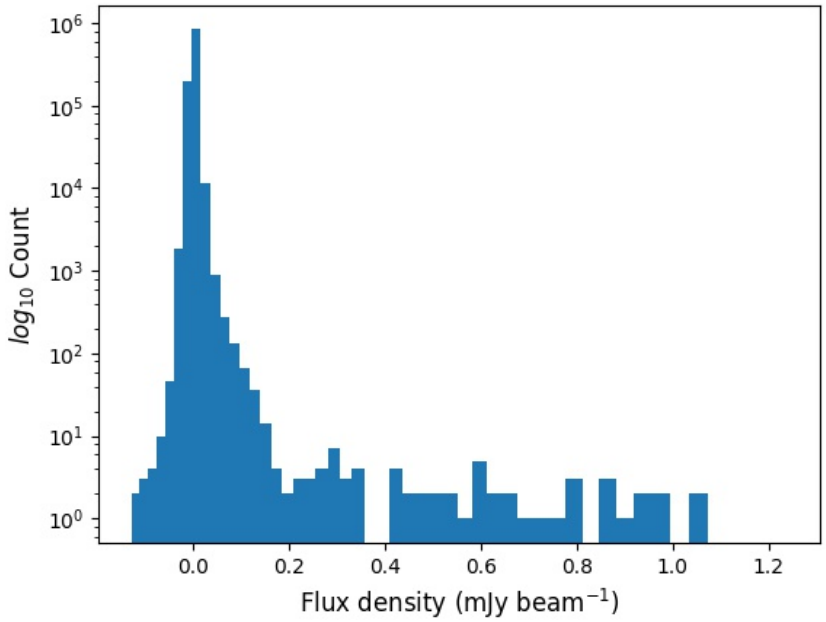


Fig. 6.4 Log histogram of the XMM image from 6.3. Only a smaller proportion of the pixels are above the noise threshold of 0.07mJybeam^{-1}.

the fields. Each facet is then subtracted and solved independently. This means each source in the field is calibrated using the nearest possible calibrator source. Applying this technique to 150MHz LOFAR data, Williams et al. (2016) have produced a 19 square degree image of the Boötes field at an order of magnitude higher resolution and sensitivity than previously achieved in this waveband. The high quality of this data serves as motivation for the work in the following section.

6.1.2 Ellipticity Measurements in the Boötes Field

As I have discussed at length there is great potential in radio observations for the future of lensing cosmology. Although there are few examples of a shear signal detection in radio in the GHz regime, current surveys do not generally have the sensitivity and resolution required to make a reliable measurement, particularity at low frequencies. Despite this, there is still some useful information to be obtained from studying ellipticity in recent radio observations. The true cosmic signal in measured correlations is likely to be smaller than the noise properties. This means a non-zero detection is a measure of the systematics and bias of the observation. This will be in part due to instrumental effects or errors in the imaging process, and also a measure of intrinsic alignment. If IA contamination is different in the

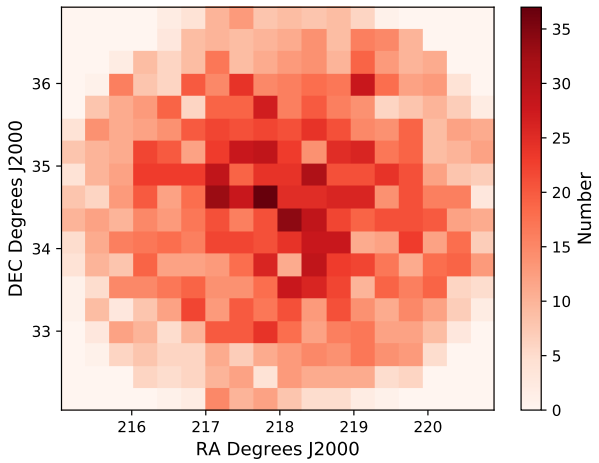


Fig. 6.5 Two-dimensional histogram of resolved sources in the Boötes field in 20 arc-minute bins.

radio population compared to optical, this may afford a strong motivation for measuring lensing at both frequencies as we may be able to isolate the cosmic signal.

The LOFAR HBA Boötes field image covers a wide area of 19 square degrees near the northern galactic pole, making it mostly free of foreground contamination. The image was created using the latest calibration and imaging techniques providing a resolution of $5.6'' \times 7.4''$ and a noise limit of ~ 0.1Jy beam^{-1}. This far exceeds any other survey at this frequency. In total 5652 sources were detected and catalogued in the field, giving a density of 0.083arcmin^{-2}. I present an initial look at the ellipticity statistics of this data and show a clear non-zero measurement in the auto-correlation. At small scales, where the cross-correlation is also non-zero, this suggests a shape bias most likely due to artefacts around bright objects in the LOFAR image. On scales larger than an arc-minute however, the cross-correlation is very small and indicates that this may be a measurable IA signal.

Of the sources detected in the Boötes field there are $3,999$ that are resolved after deconvolution from the LOFAR beam and hence suitable for shape measurement. Ellipticity and position angle were measured using PyBDSM (Mohan & Rafferty, 2015). The reader is refered to the observation paper, Williams et al. (2016), for further details. The sources have a mean absolute ellipticity of $\bar{\varepsilon} = 0.36$, with a variance of $\sigma_\varepsilon^2 = 0.04$. Figure 6.5 shows the distribution of source density across the FOV. This peaks in the centre of the image at around 0.1arcmin^{-2} dropping to ~ 0.01arcmin^{-2} toward the edge.

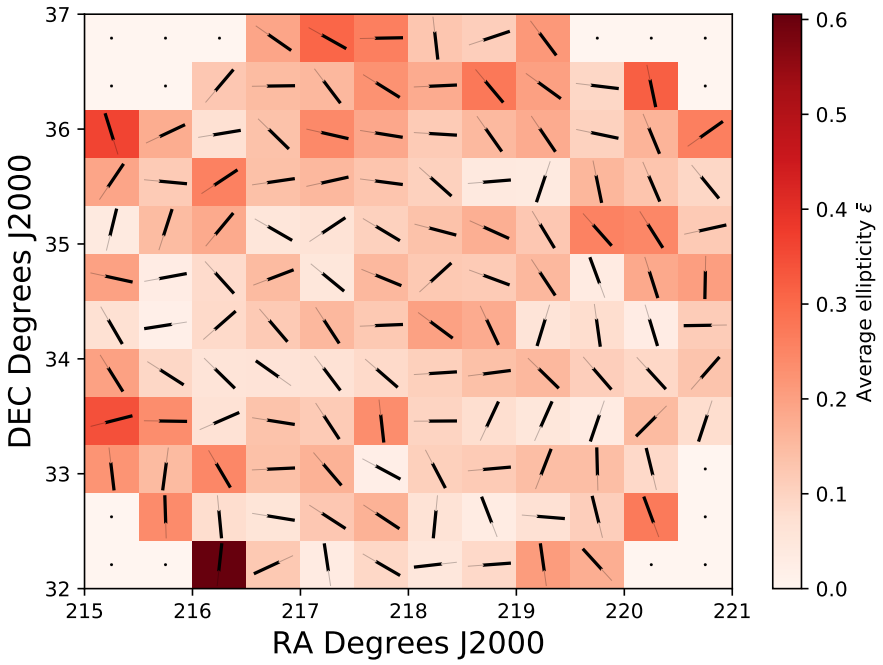


Fig. 6.6 Ellipticity averaged over 20 arc-minute pixels. Vectors show average orientation, colour map indicates $\bar{\varepsilon} = \sqrt{\bar{\varepsilon}_1^2 + \bar{\varepsilon}_2^2}$.

The average ellipticity across the field is shown in figures 6.6 and 6.7. From even these simple plots it is clear there are large scale residual systematics across the image. These averages don't vary much with either RA or DEC suggesting there could be some distortion across the image from an incomplete deconvolution. Also, notice that near the edges of the FOV the lower number densities result in a larger variation in the average ellipticity. This may result in large scales ($\sim 5^\circ$) being extremely noisy so I limit correlation calculations to smaller scales from now on.

The number of pairs at a certain separation will define the noise level for that scale. Figure 6.8 shows a count of these pairs for the relevant scales in the Boötes image. On all scales above an arc-minute there are over 10^3 pairs for bins of this size, which will provide a precise measure of the correlation function.

The galaxy-galaxy lensing statistics, $\langle \varepsilon_1^r \rangle$ and $\langle \varepsilon_2^r \rangle$, as described in section 4.5, are measures of the E-mode and B-mode lensing patterns when applied to source-background galaxy pairs. Since there should be no detectable B-mode, this acts as a measure of the systematics in any E-mode signal. As discussed, the noise properties of this data mean any true weak lensing

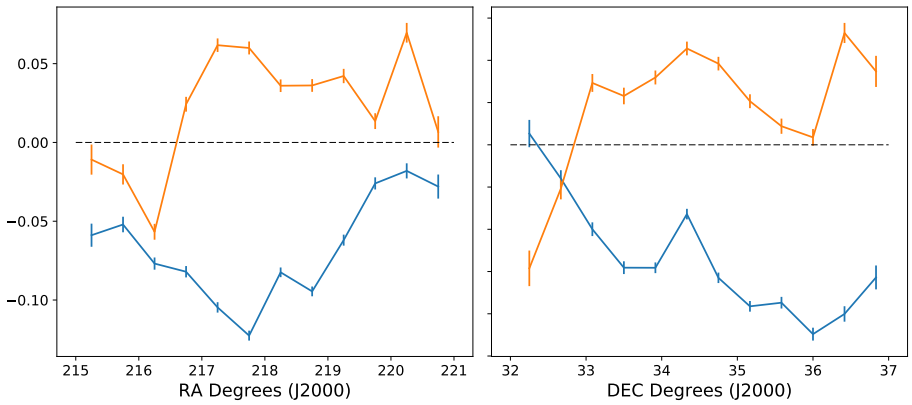


Fig. 6.7 Ellipticity measurements averaged over 20 arc-minute pixels as in figure 6.6, marginalised over declination and right ascension respectively. $\bar{\varepsilon}_1$ shown in blue, $\bar{\varepsilon}_2$ in orange. Error-bars calculated as $\sigma^2_{\gamma_i}/\sqrt{n}$.

signal is unmeasurable, but the power of E/B-mode like systematics can be characterised. Since a lensing signal is not expected around all galaxies the $\langle \varepsilon_1^r \rangle$ and $\langle \varepsilon_2^r \rangle$ statistics for every galaxy pair should average to zero. However, figure 6.9 shows a clear response at scales smaller than an arc-minute demonstrating a large lensing systematic. On these scales the B-mode signal is very similar to E-mode, supporting the theory that these systematics could be driven by ringing, or other imaging artefacts, around sources. On larger scales however, both these statistics drop to around 10^{-2}, a level at which any intrinsic alignment signal may be measured. It is therefore interesting to find that there is no appreciable alignment signal in this sample.

The auto-correlation C_{ii}, and cross-correlation C_{ij} of these statistics, as described in section 4.5 can help to determine the level of systematic noise in the data, and ultimately measure the cosmic signal. These correlations can be seen in figure 6.10. Unfortunately the autocorrelations are much higher than the 10^{-4} signal expected dues to the large scale systematics seen in the average ellipticity, and thus cannot be used to constrain cosmic shear. The cross-correlation however, shows that parity of ellipticity in the field is consistent with zero on scales over an arc-minute.

The $\xi_\pm$ correlation functions can be seen in figure 6.11. Again for ξ_+ we see that the large scale E-mode contamination completely dominates the expected 10^{-4} signal. However, ξ_- shows amplitudes of $\sim 10^{-3}$ to 10^{-4} on scales from $1'$ to $10'$ and above. This is still an order of magnitude above the results from DES Y1 data seen in figure 4.5. However, it

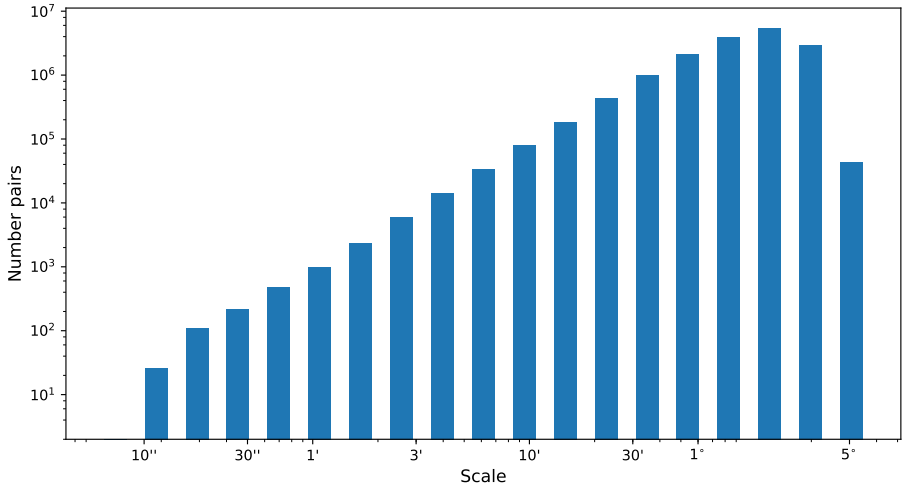


Fig. 6.8 log-log histogram of the number of galaxy pairs in the Boötes field at scales from 10” to 5°.

suggests that using a more rigorous ellipticity fitting method on this data we may detect the cosmic shear signal from this statistic.

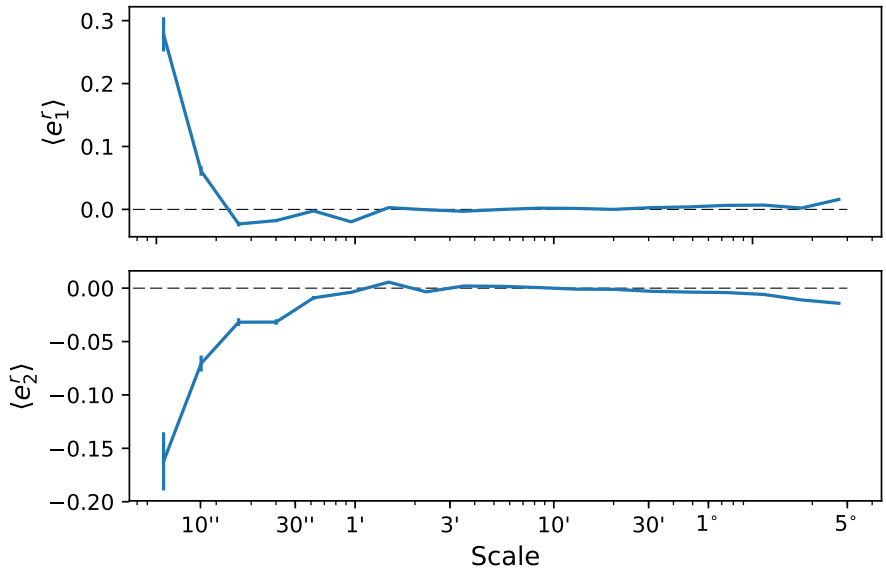


Fig. 6.9 Galaxy-galaxy lensing signal for ε_1^r (top) and ε_2^r (bottom). Error-bars show $\sigma_\varepsilon^2/\sqrt{N}$.

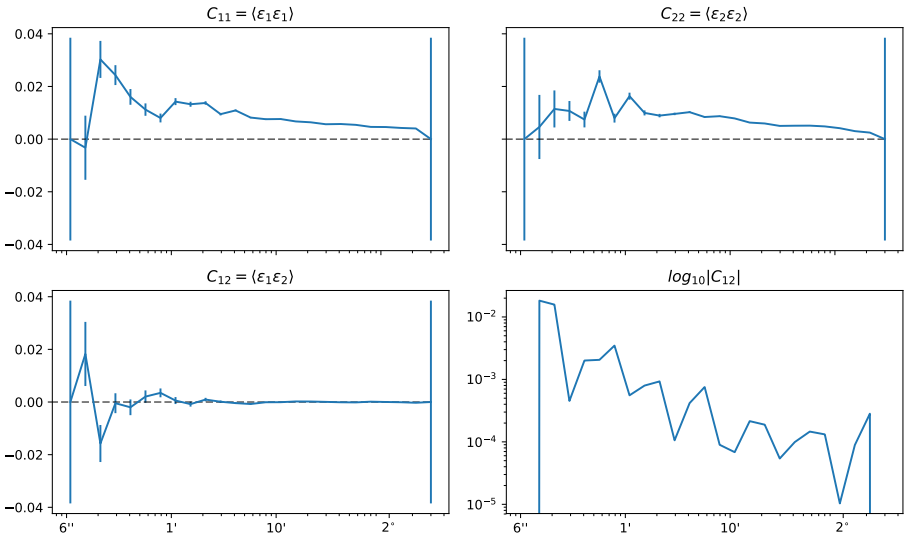


Fig. 6.10 Top left: auto-correlation of ε_1^r. Top right: auto-correlation of ε_2^r. Bottom left, right: $\varepsilon_1^r \varepsilon_2^r$ cross-correlation shown on log-linear and log-log axis respectively. Error-bars indicate $\sigma_\varepsilon^2/\sqrt{N}$.

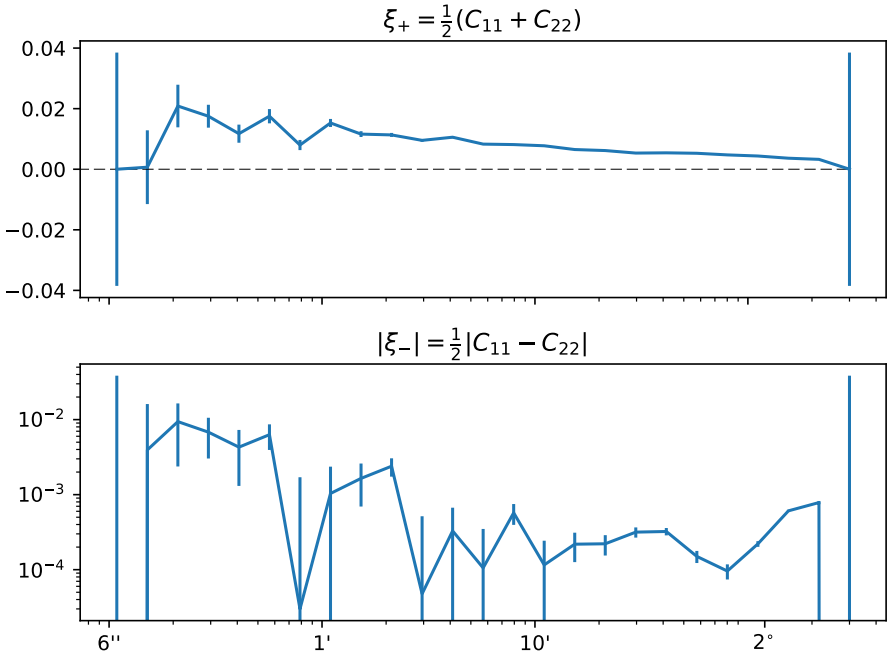


Fig. 6.11 Top: The ξ_+ correlation function of shear. Bottom: The absolute ξ_- correlation shown in Log-log. Error-bars indicate $\sigma_\varepsilon^2/\sqrt{N}$.

6.1.3 Radio-Optical Cross matching using Machine Learning

The work described in this section was conducted in collaboration with X. Morice-Atkinson and M. Withers as part of the ICG machine learning working group.

As discussed in section 3 cross matching radio and optical galaxies can provide a wealth of benefit to many science cases but is non-trivial and very time consuming. This is true even for simple single component objects due to the variation in brightness across the spectrum. The problem of matching radio sources with complex multi-component morphologies is even greater. I wanted to explore the possibility of using machine learning methods to cross match multi wavelength data, since they are ideally suited to this sort of classification task. My role in this group was to manage the project overall and provide astrophysical details of multi-band cross-matching.

We first considered the case of radio objects with single component morphologies where the sources of matching error are the lack of an observable counterpart, different source densities creating multiple false matches and astrometric error. Work by Nisbet (2016) has demonstrated that multi-band likelihoods, first developed by Sutherland & Saunders (1992), can make an efficient radio-optical matching technique. Hence, we considered a machine learning approach for constructing a similar method, using position and magnitude information.

A random forest is ideal for this problem of finding relationships between parameters, or "features", and a classification (Breiman, 2001). A random forest works by creating multiple decision trees which each split the data using the available data features (f_i), e.g. $f_3 < 0$-> $f_7 > 10$-> f_0 etc. The aim of these decisions is to minimise the "entropy" or disorder after each step. Thus, at the end of each decision tree the data should be categorised by a random set of the features. Combining all these trees and adding some non-linearity so that model is not over-fitted creates a random forest.

The first step in creating this model was to identify the data features on which the decision trees would be based. These need to be sensitive to the matching categorisation we wished to make. As mention above, the ratios of flux between objects and their position separation can be used as estimators of the match probability. So, for each optical object within $10''$ of the radio sources centroid, we chose the following as features,

- f_0: The separation between the two objects;

- f_{1-10}: The ratio of the radio source's flux at 1.4GHz with the optical object's magnitude for the five SDSS colours (u,g,r,i,z), and 5 infra-red SWIRE bands (3.6, 4.5, 5.8, 8, 24)μm;

- f_{11-20}: The ratio of the radio source's flux at 5GHz with the optical magnitude for the SDSS colours and IR SWIRE bands.

For the training set the feature f_{21} represented the match solution for that optical source, either 1 or 0. This answer could also be set to a probability which would then give a more nuanced solution, with the option of cutting matches below some threshold.

In order to develop and train this random forest we created a simulation of matches. We based this simulation on data from the Australia Telescope Large Area Survey (ATLAS) Norris et al. (2006). ATLAS is a 3.7 square degree observation near the Chandra Deep Field South and contains 726 radio sources. It is largely coincident with the infra-red Spitzer Wide-Area Extragalactic (SWIRE) Survey (Lonsdale et al., 2003), and Norris et al. (2006) present matches to 682 SWIRE sources. In addition, this area is covered by SDSS photometry. The simulation of features was then created in the following way,

1. Generate random 1.4GHz and 5GHz fluxes for the radio source drawn from a distribution in Norris et al. (2006);

2. Generate the separation and magnitudes of the matching counterpart with:

   - Position separation given by a Gaussian offset with a variance of $2''$; the typical seeing in SDSS;
   - Magnitudes randomly generated from a brightness distribution fitted to objects associated with all radio objects that have a flux within a mJy of the radio source at 1.4GHz;
   - If a magnitude falls below some observing threshold remove it for that band;

3. Calculate the number of additional candidates from a Poisson distribution with a mean of $\lambda = 2$. This provides a source ratio typically expected from current generation instruments;

4. For each additional source draw position separation from a $(0', 10')$ uniform distribution;

5. For each additional source draw magnitudes from SDSS and SWIRE distributions.

6. Calculate radio:optical and radio:IR flux ratios

We used 10^5 radio source counterparts to train the initial random forest model. Testing on a simulation of 10^6 objects a nearest neighbour match only identified the correct counterpart $\sim$65 percent of the time. Conversely, the machine learning method achieved a successful cross-match rate of over 80 percent. Clearly there is much work to do before this method could be applied to real data. However, it could aid in eliminating the bulk of cases where a match is clear.

Matching objects with complex morphologies is much more challenging. One approach for these cases considered by the group was the use of a Convolutional neural network (CNN). The method would use cut-outs of the radio intensity maps along with models of potential nearby matches in each relevant band. The input for an n$\times$n image in this model would then be an n$\times$n$\times$m data cube with the m dimension representing the various radio, optical and IR wavebands. The solutions for training could then be provided by an n$\times$n heat map representing the probability of the counterpart location. This sort of product is often referred to as "attention maps" since they are used more generally to locate items in an image. I refer the reader to Chollet et al. (2015) for further details.

As a result of this working group a new PhD project in cosmological machine learning methods has begun.

6.2 Summary and conclusions

Radio instrumentation is clearly advancing quickly and has much to offer the future of astronomy and cosmology. Radio observations have the potential to enhance almost every science case in modern astronomy but only if research keeps pace with the technological developments. There has been a dramatic increase in the global radio astronomy community in recent years as institutions seek to exploit this new and exciting regime. This thesis has aimed to build on this research and explore new techniques that will maximise the future success of radio cosmology.

In chapter 1 I outlined the big picture of cosmology providing context and background for the rest of the work. Starting with a general metric for the Universe I presented the theoretical underpinning that defines the current model of cosmology. I described problems with this model where it cannot fully explain modern observations and some of the theoretical solutions. I also outlined the various parameters that are used to describe the Universe and

constrain cosmology. Finally, I presented the probes of these parameters that have contributed to the current model and detailed the latest constraints.

In chapter 2 I gave an overview of radio astronomy and instrumentation. I showed how the transparency of the atmosphere to the radio wave band, along with the ability to combine an array of antennas is making radio observations ever more sophisticated and valuable. I discussed the Fourier relationship between interferometric radio data and sky brightness that forms an important part of my work in chapter 5. I provided the typical processes that are undertaken in order to produce radio images, mainly gridding, weighting and deconvolution, illustrating some of their undesirable effects. I concluded by reviewing the astrophysical mechanisms of radio galaxies and discussing studies of their population that are significant for the work presented in chapter 3.

Having demonstrated the theory of radio astronomy and its enormous potential benefit I then considered ambitious future projects and presented some early work in preparation for them. One such project is WEAVE-LOFAR that forms the motivation for the work in chapter 3. In this chapter, I presented the data collected from five radio surveys and the subsequent spectrograph follow-up.

Firstly, I discussed the procedure of cross-matching radio and optical galaxies and the challenge presented by complex radio morphologies. For the large percentage of compact radio objects, mostly star-forming galaxies, a simple matching technique based on separation distance is sufficient. There may be a small false positive rate when using this method, but it has been shown by McAlpine et al. (2012) that this can be reduced using multi-band likelihood matching. In the case of complex multi-component objects, which form approximately one percent of the population, we performed the cross-matching by eye by comparing optical cut-outs with radio contours to identify the most likely counterpart. Even though only a few hundred candidates were considered this approach was extremely time consuming and produced multiple conflicts between matches from different team members. This presents a significant problem for a survey on the scale on WEAVE-LOFAR particularly for science cases where these unusual sources are important. The current solution for this matching problem is an internal galaxy zoo style programme similar to the hand matching performed in this work but with a greater efficiency and number of participants. Clearly there is much work to be done in combining data from different waveband catalogues not just between radio and optical. One potential alternative is to use simulations coupled with machine learning

methods as discussed above in section 6.1.3.

Chapter 3 also contains an analysis of the quality and efficiency of the spectrograph observations. Overall, the spectroscopic reduction was reliable for objects identified as galaxies. However, for the unusually large number of QSOs in this sample the efficiency was only $\sim$50 percent and there were multiple problems with the BOSS classifications. For future surveys that will also be radio selected such as WEAVE-LOFAR the fitting procedures need to be improved for these objects. The population of this sample follows the broad trend expected from semi-analytic models such as Wilman et al. (2008) except for the star-forming sources at fluxes above 20mJy. For this sub-sample Wilman et al. (2008) seems to underestimate the normalised number counts found. This may indicate an odd feature of this area and will make an interesting study for WEAVE-LOFAR.

In chapter 4 I formulated gravitational lensing using Fermat's Principle and the parallels between light geodesics and refraction. I described the observables with which weak lensing can be measured, in particular shear (γ), and their relationship to the matter distribution of the Universe. I then considered the challenges present in optical lensing studies and proposed the alternative of radio. I discussed the various benefits and challenges of a radio approach to lensing and argued that most problems are due to the relative infancy of radio astronomy whilst its inherent advantages could make this an invaluable source of data. Concluding this chapter, I highlighted work that demonstrates the plausibility of using radio surveys for weak lensing and the potential impact of next generation instruments on cosmology.

Having detailed the potential of the radio waveband for cosmology, and the dependence of this on improved imaging or shear measurement techniques, I turned to the technical details of radio data and its relationship with shear. I begin chapter 5 by developing a mathematical model of a generalised radio observation in Fourier space. This formalism demonstrated that the circular and non-circular contributions from galaxy shapes is separable in radio data. Furthermore, the distortion of data due to ellipticity was shown to be a function of position and galaxy profile that could be modelled. From this analysis I developed the "F.I.L.M" estimator in section 5.1.2.

In order to test FILM I created a radio simulation based on the semi-empirical models of Wilman et al. (2008). I describe the simulation and subsequent analysis and testing pipeline software in section 5.2. For a basic high quality simulation I found that I could reconstruct with this accuracy thus demonstrating that in principle it is possible to measure this signal

directly in this way.

I then tested my method against simulations of increasing noise and realism. Increases in the noise from shape dispersion produce a signal with lower S/N as expected; with larger areas and more aggressive smoothing this effects can be reduced. I also tests the effect of isotropic and anisotropic sampling and noise functions. I find little degradation in the signal from the circular cases but there remains a problem with the SKA-like sampled data. We are still working on the optimal weighting to remove this effect. The contamination of the result from astrometric error and foreground sources was not significant. However, this could still be improved using better galaxy models and combining with optical catalogue positions.

In this chapter I presented three other projects related to the future of radio cosmology. First, I demonstrated some of the challenges in creating an automatic imaging pipeline for wide-field low-frequency radio data. I then look at one of the most advanced science quality images recently produced by LOFAR. I show that this image has large scale systematics remaining in the ellipticity information, but that these are at a level such that a detection of the IA signal and possible cosmic shear from ξ_- could be made. This is very promising for the next iteration of LOFAR observations. Finally, I discuss the potential of new machine learning techniques to help match the future wealth of multi data. This matching is imperative if the next generation of instruments are to reach their full potential.

In this thesis I have taken a broad look at the rapidly advancing field of radio astronomy and considered its potential impact on cosmology. I have looked specifically at the accuracy of current spectroscopic fitting methods for the upcoming LOFAR-WEAVE survey. I have also developed a novel lensing shear estimator which acts as a proof of concept for direct radio lensing measurements. Going forward the FILM method needs to be refined with a more sophisticated approach to removing the PSF contamination. I have also shown that the most recent images from LOFAR are near to producing useful cosmological results, an exciting milestone in the history of radio cosmology. This thesis has presented techniques and methods that will be essential to the next-generation potential of cosmology with radio data.

References

Abazajian, K. N., Adelman-McCarthy, J. K., Agüeros, M. A., Allam, S. S., Allende Prieto, C., An, D., Anderson, K. S. J., Anderson, S. F., Annis, J., Bahcall, N. A., & et al. (2009). The Seventh Data Release of the Sloan Digital Sky Survey. ApJS, 182, 543–558.

Abbott, B. P., Abbott, R., Abbott, T. D., Abernathy, M. R., Acernese, F., Ackley, K., Adams, C., Adams, T., Addesso, P., Adhikari, R. X., Adya, V. B., Affeldt, C., Agathos, M., Agatsuma, K., Aggarwal, N., Aguiar, O. D., Aiello, L., Ain, A., Ajith, P., Allen, B., Allocca, A., Altin, P. A., Anderson, S. B., Anderson, W. G., Arai, K., Arain, M. A., Araya, M. C., Arceneaux, C. C., Areeda, J. S., Arnaud, N., Arun, K. G., Ascenzi, S., Ashton, G., Ast, M., Aston, S. M., Astone, P., Aufmuth, P., Aulbert, C., Babak, S., Bacon, P., Bader, M. K. M., Baker, P. T., Baldaccini, F., Ballardin, G., Ballmer, S. W., Barayoga, J. C., Barclay, S. E., Barish, B. C., Barker, D., Barone, F., Barr, B., Barsotti, L., Barsuglia, M., Barta, D., Bartlett, J., Barton, M. A., Bartos, I., Bassiri, R., Basti, A., Batch, J. C., Baune, C., Bavigadda, V., Bazzan, M., Behnke, B., Bejger, M., Belczynski, C., Bell, A. S., Bell, C. J., Berger, B. K., Bergman, J., Bergmann, G., Berry, C. P. L., Bersanetti, D., Bertolini, A., Betzwieser, J., Bhagwat, S., Bhandare, R., Bilenko, I. A., Billingsley, G., Birch, J., Birney, R., Birnholtz, O., Biscans, S., Bisht, A., Bitossi, M., Biwer, C., Bizouard, M. A., Blackburn, J. K., Blair, C. D., Blair, D. G., Blair, R. M., Bloemen, S., Bock, O., Bodiya, T. P., Boer, M., Bogaert, G., Bogan, C., Bohe, A., Bojtos, P., Bond, C., Bondu, F., Bonnand, R., Boom, B. A., Bork, R., Boschi, V., Bose, S., Bouffanais, Y., Bozzi, A., Bradaschia, C., Brady, P. R., Braginsky, V. B., Branchesi, M., Brau, J. E., Briant, T., Brillet, A., Brinkmann, M., Brisson, V., Brockill, P., Brooks, A. F., Brown, D. A., Brown, D. D., Brown, N. M., Buchanan, C. C., Buikema, A., Bulik, T., Bulten, H. J., Buonanno, A., Buskulic, D., Buy, C., Byer, R. L., Cabero, M., Cadonati, L., Cagnoli, G., Cahillane, C., Bustillo, J. C., Callister, T., Calloni, E., Camp, J. B., Cannon, K. C., Cao, J., Capano, C. D., Capocasa, E., Carbognani, F., Caride, S., Diaz, J. C., Casentini, C., Caudill, S., Cavaglià, M., Cavalier, F., Cavalieri, R., Cella, G., Cepeda, C. B., Baiardi, L. C., Cerretani, G., Cesarini, E., Chakraborty, R., Chalermsongsak, T., Chamberlin, S. J., Chan, M., Chao, S., Charlton, P., Chassande-Mottin, E., Chen, H. Y., Chen, Y., Cheng, C., Chincarini, A., Chiummo, A., Cho, H. S., Cho, M., Chow, J. H., Christensen, N., Chu, Q., Chua, S., Chung, S., Ciani, G., Clara, F., Clark, J. A., Cleva, F., Coccia, E., Cohadon, P.-F., Colla, A., Collette, C. G., Cominsky, L., Constancio, M., Conte, A., Conti, L., Cook, D., Corbitt, T. R., Cornish, N., Corsi, A., Cortese, S., Costa, C. A., Coughlin, M. W., Coughlin, S. B., Coulon, J.-P., Countryman, S. T., Couvares, P., Cowan, E. E., Coward, D. M., Cowart, M. J., Coyne, D. C., Coyne, R., Craig, K., Creighton, J. D. E., Creighton, T. D., Cripe, J., Crowder, S. G., Cruise, A. M., Cumming, A., Cunningham, L., Cuoco, E., Canton, T. D., Danilishin, S. L., D'Antonio, S., Danzmann, K., Darman, N. S., Da Silva Costa, C. F., Dattilo, V., Dave, I., Daveloza, H. P., Davier, M., Davies, G. S., Daw, E. J., Day, R., De, S., DeBra, D., Debreczeni, G., Degallaix, J., De Laurentis,

M., Deléglise, S., Del Pozzo, W., Denker, T., Dent, T., Dereli, H., Dergachev, V., DeRosa, R. T., De Rosa, R., DeSalvo, R., Dhurandhar, S., Díaz, M. C., Di Fiore, L., Di Giovanni, M., Di Lieto, A., Di Pace, S., Di Palma, I., Di Virgilio, A., Dojcinoski, G., Dolique, V., Donovan, F., Dooley, K. L., Doravari, S., Douglas, R., Downes, T. P., Drago, M., Drever, R. W. P., Driggers, J. C., Du, Z., Ducrot, M., Dwyer, S. E., Edo, T. B., Edwards, M. C., Effler, A., Eggenstein, H.-B., Ehrens, P., Eichholz, J., Eikenberry, S. S., Engels, W., Essick, R. C., Etzel, T., Evans, M., Evans, T. M., Everett, R., Factourovich, M., Fafone, V., Fair, H., Fairhurst, S., Fan, X., Fang, Q., Farinon, S., Farr, B., Farr, W. M., Favata, M., Fays, M., Fehrmann, H., Fejer, M. M., Feldbaum, D., Ferrante, I., Ferreira, E. C., Ferrini, F., Fidecaro, F., Finn, L. S., Fiori, I., Fiorucci, D., Fisher, R. P., Flaminio, R., Fletcher, M., Fong, H., Fournier, J.-D., Franco, S., Frasca, S., Frasconi, F., Frede, M., Frei, Z., Freise, A., Frey, R., Frey, V., Fricke, T. T., Fritschel, P., Frolov, V. V., Fulda, P., Fyffe, M., Gabbard, H. A. G., Gair, J. R., Gammaitoni, L., Gaonkar, S. G., Garufi, F., Gatto, A., Gaur, G., Gehrels, N., Gemme, G., Gendre, B., Genin, E., Gennai, A., George, J., Gergely, L., Germain, V., Ghosh, A., Ghosh, A., Ghosh, S., Giaime, J. A., Giardina, K. D., Giazotto, A., Gill, K., Glaefke, A., Gleason, J. R., Goetz, E., Goetz, R., Gondan, L., González, G., Castro, J. M. G., Gopakumar, A., Gordon, N. A., Gorodetsky, M. L., Gossan, S. E., Gosselin, M., Gouaty, R., Graef, C., Graff, P. B., Granata, M., Grant, A., Gras, S., Gray, C., Greco, G., Green, A. C., Greenhalgh, R. J. S., Groot, P., Grote, H., Grunewald, S., Guidi, G. M., Guo, X., Gupta, A., Gupta, M. K., Gushwa, K. E., Gustafson, E. K., Gustafson, R., Hacker, J. J., Hall, B. R., Hall, E. D., Hammond, G., H (2016). Observation of gravitational waves from a binary black hole merger. *Phys. Rev. Lett.*, 116, 061102.

Abramovici, A., Althouse, W. E., Drever, R. W. P., Gursel, Y., Kawamura, S., Raab, F. J., Shoemaker, D., Sievers, L., Spero, R. E., & Thorne, K. S. (1992). LIGO - The Laser Interferometer Gravitational-Wave Observatory. *Science*, 256, 325–333.

Adelman-McCarthy, J. K., Agüeros, M. A., Allam, S. S., Allende Prieto, C., Anderson, K. S. J., Anderson, S. F., Annis, J., Bahcall, N. A., Bailer-Jones, C. A. L., Baldry, I. K., Barentine, J. C., Bassett, B. A., Becker, A. C., Beers, T. C., Bell, E. F., Berlind, A. A., Bernardi, M., Blanton, M. R., Bochanski, J. J., Boroski, W. N., Brinchmann, J., Brinkmann, J., Brunner, R. J., Budavári, T., Carliles, S., Carr, M. A., Castander, F. J., Cinabro, D., Cool, R. J., Covey, K. R., Csabai, I., Cunha, C. E., Davenport, J. R. A., Dilday, B., Doi, M., Eisenstein, D. J., Evans, M. L., Fan, X., Finkbeiner, D. P., Friedman, S. D., Frieman, J. A., Fukugita, M., Gänsicke, B. T., Gates, E., Gillespie, B., Glazebrook, K., Gray, J., Grebel, E. K., Gunn, J. E., Gurbani, V. K., Hall, P. B., Harding, P., Harvanek, M., Hawley, S. L., Hayes, J., Heckman, T. M., Hendry, J. S., Hindsley, R. B., Hirata, C. M., Hogan, C. J., Hogg, D. W., Hyde, J. B., Ichikawa, S.-i., Ivezić, Ž., Jester, S., Johnson, J. A., Jorgensen, A. M., Jurić, M., Kent, S. M., Kessler, R., Kleinman, S. J., Knapp, G. R., Kron, R. G., Krzesinski, J., Kuropatkin, N., Lamb, D. Q., Lampeitl, H., Lebedeva, S., Lee, Y. S., French Leger, R., Lépine, S., Lima, M., Lin, H., Long, D. C., Loomis, C. P., Loveday, J., Lupton, R. H., Malanushenko, O., Malanushenko, V., Mandelbaum, R., Margon, B., Marriner, J. P., Martínez-Delgado, D., Matsubara, T., McGehee, P. M., McKay, T. A., Meiksin, A., Morrison, H. L., Munn, J. A., Nakajima, R., Neilsen, Jr., E. H., Newberg, H. J., Nichol, R. C., Nicinski, T., Nieto-Santisteban, M., Nitta, A., Okamura, S., Owen, R., Oyaizu, H., Padmanabhan, N., Pan, K., Park, C., Peoples, Jr., J., Pier, J. R., Pope, A. C., Purger, N., Raddick, M. J., Re Fiorentin, P., Richards, G. T., Richmond, M. W., Riess, A. G., Rix, H.-W., Rockosi, C. M., Sako, M., Schlegel, D. J., Schneider, D. P., Schreiber, M. R., Schwope, A. D., Seljak, U., Sesar, B., Sheldon, E., Shimasaku, K., Sivarani, T., Allyn

Smith, J., Snedden, S. A., Steinmetz, M., Strauss, M. A., SubbaRao, M., Suto, Y., Szalay, A. S., Szapudi, I., Szkody, P., Tegmark, M., Thakar, A. R., Tremonti, C. A., Tucker, D. L., Uomoto, A., Vanden Berk, D. E., Vandenberg, J., Vidrih, S., Vogeley, M. S., Voges, W., Vogt, N. P., Wadadekar, Y., Weinberg, D. H., West, A. A., White, S. D. M., Wilhite, B. C., Yanny, B., Yocum, D. R., York, D. G., Zehavi, I., & Zucker, D. B. (2008). The Sixth Data Release of the Sloan Digital Sky Survey. ApJS, 175, 297–313.

Ahn, C. P., Alexandroff, R., Allende Prieto, C., Anderson, S. F., Anderton, T., Andrews, B. H., Aubourg, É., Bailey, S., Balbinot, E., Barnes, R., & et al. (2012). The Ninth Data Release of the Sloan Digital Sky Survey: First Spectroscopic Data from the SDSS-III Baryon Oscillation Spectroscopic Survey. ApJS, 203, 21.

Alam, S., Albareti, F. D., Allende Prieto, C., Anders, F., Anderson, S. F., Anderton, T., Andrews, B. H., Armengaud, E., Aubourg, É., Bailey, S., & et al. (2015). The Eleventh and Twelfth Data Releases of the Sloan Digital Sky Survey: Final Data from SDSS-III. ApJS, 219, 12.

Alam, S., Ata, M., Bailey, S., Beutler, F., Bizyaev, D., Blazek, J. A., Bolton, A. S., Brownstein, J. R., Burden, A., Chuang, C.-H., Comparat, J., Cuesta, A. J., Dawson, K. S., Eisenstein, D. J., Escoffier, S., Gil-Marín, H., Grieb, J. N., Hand, N., Ho, S., Kinemuchi, K., Kirkby, D., Kitaura, F., Malanushenko, E., Malanushenko, V., Maraston, C., McBride, C. K., Nichol, R. C., Olmstead, M. D., Oravetz, D., Padmanabhan, N., Palanque-Delabrouille, N., Pan, K., Pellejero-Ibanez, M., Percival, W. J., Petitjean, P., Prada, F., Price-Whelan, A. M., Reid, B. A., Rodríguez-Torres, S. A., Roe, N. A., Ross, A. J., Ross, N. P., Rossi, G., Rubiño-Martín, J. A., Sánchez, A. G., Saito, S., Salazar-Albornoz, S., Samushia, L., Satpathy, S., Scóccola, C. G., Schlegel, D. J., Schneider, D. P., Seo, H.-J., Simmons, A., Slosar, A., Strauss, M. A., Swanson, M. E. C., Thomas, D., Tinker, J. L., Tojeiro, R., Vargas Magaña, M., Vazquez, J. A., Verde, L., Wake, D. A., Wang, Y., Weinberg, D. H., White, M., Wood-Vasey, W. M., Yèche, C., Zehavi, I., Zhai, Z., & Zhao, G.-B. (2016). The clustering of galaxies in the completed SDSS-III Baryon Oscillation Spectroscopic Survey: cosmological analysis of the DR12 galaxy sample. *ArXiv e-prints*.

Alcock, C., Allsman, R. A., Alves, D. R., Axelrod, T. S., Becker, A. C., Bennett, D. P., Cook, K. H., Dalal, N., Drake, A. J., Freeman, K. C., Geha, M., Griest, K., Lehner, M. J., Marshall, S. L., Minniti, D., Nelson, C. A., Peterson, B. A., Popowski, P., Pratt, M. R., Quinn, P. J., Stubbs, C. W., Sutherland, W., Tomaney, A. B., Vandehei, T., & Welch, D. (2000). The MACHO Project: Microlensing Results from 5.7 Years of Large Magellanic Cloud Observations. ApJ, 542, 281–307.

Alpher, R. A., Bethe, H., & Gamow, G. (1948). The origin of chemical elements. *Phys. Rev.*, 73, 803–804.

Amendola, L. & Tsujikawa, S. (2010). Cambridge, first edition.

Angulo, R. E., Baugh, C. M., Frenk, C. S., & Lacey, C. G. (2008). The detectability of baryonic acoustic oscillations in future galaxy surveys. MNRAS, 383, 755–776.

Arkani-Hamed, N., Finkbeiner, D. P., Slatyer, T. R., & Weiner, N. (2009). A theory of dark matter. Phys. Rev. D, 79(1), 015014.

Arnalte-Mur, P., Martínez, V. J., Norberg, P., Fernández-Soto, A., Ascaso, B., Merson, A. I., Aguerri, J. A. L., Castander, F. J., Hurtado-Gil, L., López-Sanjuan, C., Molino, A., Montero-Dorta, A. D., Stefanon, M., Alfaro, E., Aparicio-Villegas, T., Benítez, N., Broadhurst, T., Cabrera-Caño, J., Cepa, J., Cerviño, M., Cristóbal-Hornillos, D., del Olmo, A., González Delgado, R. M., Husillos, C., Infante, L., Márquez, I., Masegosa, J., Moles, M., Perea, J., Pović, M., Prada, F., & Quintana, J. M. (2014). The ALHAMBRA survey: evolution of galaxy clustering since z$\sim$1. MNRAS, 441, 1783–1801.

Baddour, N. (2009). Operational and convolution properties of two dimensional fourier transforms in polar coordinates. 26, 1767–1777.

Baldwin, J. A., Phillips, M., & Terlevich, R. (1981). Classification parameters for the emission-line spectra of extragalactic objects. *Publications of the Astronomical Society of the Pacific*, 93, 5–19.

Battye, R. A. & Browne, I. W. A. (2009). Radio and optical orientations of galaxies. MNRAS, 399, 1888–1900.

Beck, R. (2007). Measurements of Cosmic Magnetism with LOFAR and SKA. *Advances in Radio Science*, 5, 399–405.

Becker, R. H., White, R. L., & Edwards, A. L. (1991). A new catalog of 53,522 4.85 GHz sources. ApJS, 75, 1–229.

Becker, R. H., White, R. L., & Helfand, D. J. (1995). The FIRST Survey: Faint Images of the Radio Sky at Twenty Centimeters. ApJ, 450, 559.

Benn, C. R., Rowan-Robinson, M., McMahon, R. G., Broadhurst, T. J., & Lawrence, A. (1993). Spectroscopy of faint radio sources - The nature of the sub-mJy radio-source population. MNRAS, 263, 98–122.

Bennett, D. P., Udalski, A., Han, C., Bond, I. A., Beaulieu, J.-P., Skowron, J., Gaudi, B. S., Koshimoto, N., Abe, F., Asakura, Y., Barry, R. K., Bhattacharya, A., Donachie, M., Evans, P., Fukui, A., Hirao, Y., Itow, Y., Li, M. C. A., Ling, C. H., Masuda, K., Matsubara, Y., Muraki, Y., Nagakane, M., Ohnishi, K., Oyokawa, H., Ranc, C., Rattenbury, N. J., Rosenthal, M. K., Saito, T., Sharan, A., Sullivan, D. J., Sumi, T., Suzuki, D., Tristram, P. J., Yonehara, A., Szymanski, M. K., Poleski, R., Soszynski, I., Ulaczyk, K., Wyrzykowski, L., DePoy, D., Gould, A., Pogge, R. W., Yee, J. C., Albrow, M. D., Bachelet, E., Batista, V., Bowens-Rubin, R., Brillant, S., Caldwell, J. A. R., Cole, A., Coutures, C., Dieters, S., Dominis Prester, D., Donatowicz, J., Fouque, P., Horne, K., Hundertmark, M., Kains, N., Kane, S. R., Marquette, J.-B., Menzies, J., Pollard, K. R., Ranc, C., Sahu, K. C., Wambsganss, J., Williams, A., & Zub, M. (2017). The First Planetary Microlensing Event with Two Microlensed Source Stars. *ArXiv e-prints*.

Bertin, E. & Arnouts, S. (1996). SExtractor: Software for source extraction. A&AS, 117, 393–404.

Best, P. N., Kauffmann, G., Heckman, T. M., Brinchmann, J., Charlot, S., Ivezić, Ž., & White, S. D. M. (2005). The host galaxies of radio-loud active galactic nuclei: mass dependences, gas cooling and active galactic nuclei feedback. MNRAS, 362, 25–40.

Betoule, M., Kessler, R., Guy, J., Mosher, J., Hardin, D., Biswas, R., Astier, P., El-Hage, P., Konig, M., Kuhlmann, S., Marriner, J., Pain, R., Regnault, N., Balland, C., Bassett, B. A., Brown, P. J., Campbell, H., Carlberg, R. G., Cellier-Holzem, F., Cinabro, D., Conley, A., D'Andrea, C. B., DePoy, D. L., Doi, M., Ellis, R. S., Fabbro, S., Filippenko, A. V., Foley, R. J., Frieman, J. A., Fouchez, D., Galbany, L., Goobar, A., Gupta, R. R., Hill, G. J., Hlozek, R., Hogan, C. J., Hook, I. M., Howell, D. A., Jha, S. W., Le Guillou, L., Leloudas, G., Lidman, C., Marshall, J. L., Möller, A., Mourão, A. M., Neveu, J., Nichol, R., Olmstead, M. D., Palanque-Delabrouille, N., Perlmutter, S., Prieto, J. L., Pritchet, C. J., Richmond, M., Riess, A. G., Ruhlmann-Kleider, V., Sako, M., Schahmaneche, K., Schneider, D. P., Smith, M., Sollerman, J., Sullivan, M., Walton, N. A., & Wheeler, C. J. (2014). Improved cosmological constraints from a joint analysis of the SDSS-II and SNLS supernova samples. A&A, 568, A22.

Bianchi, L. & GALEX Team (1999). The Galaxy Evolution Explorer (GALEX): an All Sky Ultraviolet Survey. Mem. Soc. Astron. Italiana, 70.

Blake, C., Ferreira, P. G., & Borrill, J. (2004). The angular power spectrum of NVSS radio galaxies. MNRAS, 351, 923–934.

Bolton, A. S., Schlegel, D. J., Aubourg, É., Bailey, S., Bhardwaj, V., Brownstein, J. R., Burles, S., Chen, Y.-M., Dawson, K., Eisenstein, D. J., Gunn, J. E., Knapp, G. R., Loomis, C. P., Lupton, R. H., Maraston, C., Muna, D., Myers, A. D., Olmstead, M. D., Padmanabhan, N., Pâris, I., Percival, W. J., Petitjean, P., Rockosi, C. M., Ross, N. P., Schneider, D. P., Shu, Y., Strauss, M. A., Thomas, D., Tremonti, C. A., Wake, D. A., Weaver, B. A., & Wood-Vasey, W. M. (2012). Spectral Classification and Redshift Measurement for the SDSS-III Baryon Oscillation Spectroscopic Survey. AJ, 144, 144.

Bonaldi, A., Harrison, I., Camera, S., & Brown, M. L. (2016). SKA weak lensing- II. Simulated performance and survey design considerations. MNRAS, 463, 3686–3698.

Bond, J. R. & Efstathiou, G. (1984). Cosmic background radiation anisotropies in universes dominated by nonbaryonic dark matter. ApJ, 285, L45–L48.

Bondi, M., Ciliegi, P., Zamorani, G., Gregorini, L., Vettolani, G., Parma, P., de Ruiter, H., Le Fevre, O., Arnaboldi, M., Guzzo, L., Maccagni, D., Scaramella, R., Adami, C., Bardelli, S., Bolzonella, M., Bottini, D., Cappi, A., Foucaud, S., Franzetti, P., Garilli, B., Gwyn, S., Ilbert, O., Iovino, A., Le Brun, V., Marano, B., Marinoni, C., McCracken, H. J., Meneux, B., Pollo, A., Pozzetti, L., Radovich, M., Ripepi, V., Rizzo, D., Scodeggio, M., Tresse, L., Zanichelli, A., & Zucca, E. (2003). The VLA-VIRMOS Deep Field. I. Radio observations probing the mu Jy source population. A&A, 403, 857–867.

Bonvin, V., Courbin, F., Suyu, S. H., Marshall, P. J., Rusu, C. E., Sluse, D., Tewes, M., Wong, K. C., Collett, T., Fassnacht, C. D., Treu, T., Auger, M. W., Hilbert, S., Koopmans, L. V. E., Meylan, G., Rumbaugh, N., Sonnenfeld, A., & Spiniello, C. (2017). H0LiCOW - V. New COSMOGRAIL time delays of HE 0435-1223: H_0 to 3.8 per cent precision from strong lensing in a flat ΛCDM model. MNRAS, 465, 4914–4930.

Booth, R. S., de Blok, W. J. G., Jonas, J. L., & Fanaroff, B. (2009). MeerKAT Key Project Science, Specifications, and Proposals. *ArXiv e-prints*.

Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5–32.

Bridle, S., Shawe-Taylor, J., Amara, A., Applegate, D., Balan, Berge, J. S. T., Bernstein, G., Dahle, H., Erben, T., Gill, M., Heavens, A., Heymans, C., High, F. W., Hoekstra, H., Jarvis, M., Kirk, D., Kitching, T., Kneib, J.-P., Kuijken, K., Lagatutta, D., Mandelbaum, R., Massey, R., Mellier, Y., Moghaddam, B., Moudden, Y., Nakajima, R., Paulin-Henriksson, S., Pires, S., Rassat, A., Refregier, A., Rhodes, J., Schrabback, T., Semboloni, E., Shmakova, M., van Waerbeke, L., Witherick, D., Voigt, L., & Wittman, D. (2009). Handbook for the GREAT08 Challenge: An image analysis competition for cosmological lensing. *Annals of Applied Statistics*, 3, 6–37.

Briggs, D. S. (1995). High Fidelity Interferometric Imaging: Robust Weighting and NNLS Deconvolution. In *American Astronomical Society Meeting Abstracts*, volume 27 of *Bulletin of the American Astronomical Society* (pp. 1444).

Brown, M. L. & Battye, R. A. (2011). Polarization as an indicator of intrinsic alignment in radio weak lensing. MNRAS, 410, 2057–2074.

Bundy, K., Bershady, M. A., Law, D. R., Yan, R., Drory, N., MacDonald, N., Wake, D. A., Cherinka, B., Sánchez-Gallego, J. R., Weijmans, A.-M., Thomas, D., Tremonti, C., Masters, K., Coccato, L., Diamond-Stanic, A. M., Aragón-Salamanca, A., Avila-Reese, V., Badenes, C., Falcón-Barroso, J., Belfiore, F., Bizyaev, D., Blanc, G. A., Bland-Hawthorn, J., Blanton, M. R., Brownstein, J. R., Byler, N., Cappellari, M., Conroy, C., Dutton, A. A., Emsellem, E., Etherington, J., Frinchaboy, P. M., Fu, H., Gunn, J. E., Harding, P., Johnston, E. J., Kauffmann, G., Kinemuchi, K., Klaene, M. A., Knapen, J. H., Leauthaud, A., Li, C., Lin, L., Maiolino, R., Malanushenko, V., Malanushenko, E., Mao, S., Maraston, C., McDermid, R. M., Merrifield, M. R., Nichol, R. C., Oravetz, D., Pan, K., Parejko, J. K., Sanchez, S. F., Schlegel, D., Simmons, A., Steele, O., Steinmetz, M., Thanjavur, K., Thompson, B. A., Tinker, J. L., van den Bosch, R. C. E., Westfall, K. B., Wilkinson, D., Wright, S., Xiao, T., & Zhang, K. (2015). Overview of the SDSS-IV MaNGA Survey: Mapping nearby Galaxies at Apache Point Observatory. ApJ, 798, 7.

Burden, A., Percival, W. J., Manera, M., Cuesta, A. J., Vargas Magana, M., & Ho, S. (2014). Efficient reconstruction of linear baryon acoustic oscillations in galaxy surveys. *Monthly Notices of the Royal Astronomical Society*, 445(3), 3152.

Burke, B. F. & Graham-Smith, F. (2010). *An Introduction to Radio Astronomy*. Cambridge.

Burrage, C. & Sakstein, J. (2017). Tests of Chameleon Gravity. *ArXiv e-prints*.

Camera, S., Harrison, I., Bonaldi, A., & Brown, M. L. (2017). SKA weak lensing - III. Added value of multiwavelength synergies for the mitigation of systematics. MNRAS, 464, 4747–4760.

Campbell, H., D'Andrea, C. B., Nichol, R. C., Sako, M., Smith, M., Lampeitl, H., Olmstead, M. D., Bassett, B., Biswas, R., Brown, P., Cinabro, D., Dawson, K. S., Dilday, B., Foley, R. J., Frieman, J. A., Garnavich, P., Hlozek, R., Jha, S. W., Kuhlmann, S., Kunz, M., Marriner, J., Miquel, R., Richmond, M., Riess, A., Schneider, D. P., Sollerman, J., Taylor, M., & Zhao, G.-B. (2013). Cosmology with Photometrically Classified Type Ia Supernovae from the SDSS-II Supernova Survey. ApJ, 763, 88.

Capozzi, D., Maraston, C., Daddi, E., Renzini, A., Strazzullo, V., & Gobat, R. (2016). Revisiting the role of the thermally pulsating asymptotic-giant-branch phase in high-redshift galaxies. MNRAS, 456, 790–830.

Chandrasekhar, S. (1931). The Maximum Mass of Ideal White Dwarfs. ApJ, 74, 81.

Chang, T.-C. & Refregier, A. (2002). Shape Reconstruction and Weak Lensing Measurement with Interferometers: A Shapelet Approach. ApJ, 570, 447–456.

Chang, T.-C., Refregier, A., & Helfand, D. J. (2004). Weak Lensing by Large-Scale Structure with the FIRST Radio Survey. ApJ, 617, 794–810.

Cheng, T.-P. (2005). *Relativity, Gravitation and Cosmology: A Basic Introduction*. Oxford University Press.

Chollet, F. et al. (2015). Keras. https://github.com/fchollet/keras.

Ciliegi, P., McMahon, R. G., Miley, G., Gruppioni, C., Rowan-Robinson, M., Cesarsky, C., Danese, L., Franceschini, A., Genzel, R., Lawrence, A., Lemke, D., Oliver, S., Puget, J.-L., & Rocca-Volmerange, B. (1999). A deep vla survey at 20 cm of the iso elais survey regions. *Monthly Notices of the Royal Astronomical Society*, 302(2), 222–244.

Clowe, D., Bradač, M., Gonzalez, A. H., Markevitch, M., Randall, S. W., Jones, C., & Zaritsky, D. (2006). A Direct Empirical Proof of the Existence of Dark Matter. ApJ, 648, L109–L113.

Cohen, A. S., Lane, W. M., Cotton, W. D., Kassim, N. E., Lazio, T. J. W., Perley, R. A., Condon, J. J., & Erickson, W. C. (2007). The VLA Low-Frequency Sky Survey. AJ, 134, 1245–1262.

Cohen, M. H. & Shaffer, D. B. (1971). Positions of Radio Sources from Long-Baseline Interferometry. AJ, 76, 91.

Colless, M., Dalton, G., Maddox, S., Sutherland, W., Norberg, P., Cole, S., Bland-Hawthorn, J., Bridges, T., Cannon, R., Collins, C., Couch, W., Cross, N., Deeley, K., De Propris, R., Driver, S. P., Efstathiou, G., Ellis, R. S., Frenk, C. S., Glazebrook, K., Jackson, C., Lahav, O., Lewis, I., Lumsden, S., Madgwick, D., Peacock, J. A., Peterson, B. A., Price, I., Seaborne, M., & Taylor, K. (2001). The 2dF Galaxy Redshift Survey: spectra and redshifts. MNRAS, 328, 1039–1063.

Collett, T. E. (2015). The Population of Galaxy-Galaxy Strong Lenses in Forthcoming Optical Imaging Surveys. ApJ, 811, 20.

Condon, J. J. (1992). Radio emission from normal galaxies. ARA&A, 30, 575–611.

Condon, J. J., Cotton, W. D., Greisen, E. W., Yin, Q. F., Perley, R. A., Taylor, G. B., & Broderick, J. J. (1998). The NRAO VLA Sky Survey. AJ, 115, 1693–1716.

Cornwell, T., Braun, R., & Briggs, D. S. (1999). Deconvolution. In G. B. Taylor, C. L. Carilli, & R. A. Perley (Eds.), *Synthesis Imaging in Radio Astronomy II*, volume 180 of *Astronomical Society of the Pacific Conference Series* (pp. 151).

Covi, L., Kim, H. B., Kim, J. E., & Roszkowski, L. (2001). Axinos as dark matter. *Journal of High Energy Physics*, 5, 033.

Dawson, K. S., Kneib, J.-P., Percival, W. J., Alam, S., Albareti, F. D., Anderson, S. F., Armengaud, E., Aubourg, É., Bailey, S., Bautista, J. E., Berlind, A. A., Bershady, M. A., Beutler, F., Bizyaev, D., Blanton, M. R., Blomqvist, M., Bolton, A. S., Bovy, J., Brandt, W. N., Brinkmann, J., Brownstein, J. R., Burtin, E., Busca, N. G., Cai, Z., Chuang, C.-H., Clerc, N., Comparat, J., Cope, F., Croft, R. A. C., Cruz-Gonzalez, I., da Costa, L. N., Cousinou, M.-C., Darling, J., de la Macorra, A., de la Torre, S., Delubac, T., du Mas des Bourboux, H., Dwelly, T., Ealet, A., Eisenstein, D. J., Eracleous, M., Escoffier, S., Fan, X., Finoguenov, A., Font-Ribera, A., Frinchaboy, P., Gaulme, P., Georgakakis, A., Green, P., Guo, H., Guy, J., Ho, S., Holder, D., Huehnerhoff, J., Hutchinson, T., Jing, Y., Jullo, E., Kamble, V., Kinemuchi, K., Kirkby, D., Kitaura, F.-S., Klaene, M. A., Laher, R. R., Lang, D., Laurent, P., Le Goff, J.-M., Li, C., Liang, Y., Lima, M., Lin, Q., Lin, W., Lin, Y.-T., Long, D. C., Lundgren, B., MacDonald, N., Geimba Maia, M. A., Malanushenko, E., Malanushenko, V., Mariappan, V., McBride, C. K., McGreer, I. D., Ménard, B., Merloni, A., Meza, A., Montero-Dorta, A. D., Muna, D., Myers, A. D., Nandra, K., Naugle, T., Newman, J. A., Noterdaeme, P., Nugent, P., Ogando, R., Olmstead, M. D., Oravetz, A., Oravetz, D. J., Padmanabhan, N., Palanque-Delabrouille, N., Pan, K., Parejko, J. K., Pâris, I., Peacock, J. A., Petitjean, P., Pieri, M. M., Pisani, A., Prada, F., Prakash, A., Raichoor, A., Reid, B., Rich, J., Ridl, J., Rodriguez-Torres, S., Carnero Rosell, A., Ross, A. J., Rossi, G., Ruan, J., Salvato, M., Sayres, C., Schneider, D. P., Schlegel, D. J., Seljak, U., Seo, H.-J., Sesar, B., Shandera, S., Shu, Y., Slosar, A., Sobreira, F., Streblyanska, A., Suzuki, N., Taylor, D., Tao, C., Tinker, J. L., Tojeiro, R., Vargas-Magaña, M., Wang, Y., Weaver, B. A., Weinberg, D. H., White, M., Wood-Vasey, W. M., Yeche, C., Zhai, Z., Zhao, C., Zhao, G.-b., Zheng, Z., Ben Zhu, G., & Zou, H. (2016). The SDSS-IV Extended Baryon Oscillation Spectroscopic Survey: Overview and Early Data. AJ, 151, 44.

Dawson, K. S., Schlegel, D. J., Ahn, C. P., Anderson, S. F., Aubourg, É., Bailey, S., Barkhouser, R. H., Bautista, J. E., Beifiori, A., Berlind, A. A., Bhardwaj, V., Bizyaev, D., Blake, C. H., Blanton, M. R., Blomqvist, M., Bolton, A. S., Borde, A., Bovy, J., Brandt, W. N., Brewington, H., Brinkmann, J., Brown, P. J., Brownstein, J. R., Bundy, K., Busca, N. G., Carithers, W., Carnero, A. R., Carr, M. A., Chen, Y., Comparat, J., Connolly, N., Cope, F., Croft, R. A. C., Cuesta, A. J., da Costa, L. N., Davenport, J. R. A., Delubac, T., de Putter, R., Dhital, S., Ealet, A., Ebelke, G. L., Eisenstein, D. J., Escoffier, S., Fan, X., Filiz Ak, N., Finley, H., Font-Ribera, A., Génova-Santos, R., Gunn, J. E., Guo, H., Haggard, D., Hall, P. B., Hamilton, J.-C., Harris, B., Harris, D. W., Ho, S., Hogg, D. W., Holder, D., Honscheid, K., Huehnerhoff, J., Jordan, B., Jordan, W. P., Kauffmann, G., Kazin, E. A., Kirkby, D., Klaene, M. A., Kneib, J.-P., Le Goff, J.-M., Lee, K.-G., Long, D. C., Loomis, C. P., Lundgren, B., Lupton, R. H., Maia, M. A. G., Makler, M., Malanushenko, E., Malanushenko, V., Mandelbaum, R., Manera, M., Maraston, C., Margala, D., Masters, K. L., McBride, C. K., McDonald, P., McGreer, I. D., McMahon, R. G., Mena, O., Miralda-Escudé, J., Montero-Dorta, A. D., Montesano, F., Muna, D., Myers, A. D., Naugle, T., Nichol, R. C., Noterdaeme, P., Nuza, S. E., Olmstead, M. D., Oravetz, A., Oravetz, D. J., Owen, R., Padmanabhan, N., Palanque-Delabrouille, N., Pan, K., Parejko, J. K., Pâris, I., Percival, W. J., Pérez-Fournon, I., Pérez-Ràfols, I., Petitjean, P., Pfaffenberger, R., Pforr, J., Pieri, M. M., Prada, F., Price-Whelan, A. M., Raddick, M. J., Rebolo, R., Rich, J., Richards, G. T., Rockosi, C. M., Roe, N. A., Ross, A. J., Ross, N. P., Rossi, G., Rubiño-Martin, J. A., Samushia, L., Sánchez, A. G., Sayres, C.,

Schmidt, S. J., Schneider, D. P., Scóccola, C. G., Seo, H.-J., Shelden, A., Sheldon, E., Shen, Y., Shu, Y., Slosar, A., Smee, S. A., Snedden, S. A., Stauffer, F., Steele, O., Strauss, M. A., Streblyanska, A., Suzuki, N., Swanson, M. E. C., Tal, T., Tanaka, M., Thomas, D., Tinker, J. L., Tojeiro, R., Tremonti, C. A., Vargas Magaña, M., Verde, L., Viel, M., Wake, D. A., Watson, M., Weaver, B. A., Weinberg, D. H., Weiner, B. J., West, A. A., White, M., Wood-Vasey, W. M., Yeche, C., Zehavi, I., Zhao, G.-B., & Zheng, Z. (2013). The Baryon Oscillation Spectroscopic Survey of SDSS-III. AJ, 145, 10.

De Young, D. S. (1989). Star formation in radio galaxies at large redshift. ApJ, 342, L59–L62.

DES Collaboration, Abbott, T. M. C., Abdalla, F. B., Alarcon, A., Aleksić, J., Allam, S., Allen, S., Amara, A., Annis, J., Asorey, J., Avila, S., Bacon, D., Balbinot, E., Banerji, M., Banik, N., Barkhouse, W., Baumer, M., Baxter, E., Bechtol, K., Becker, M. R., Benoit-Lévy, A., Benson, B. A., Bernstein, G. M., Bertin, E., Blazek, J., Bridle, S. L., Brooks, D., Brout, D., Buckley-Geer, E., Burke, D. L., Busha, M. T., Capozzi, D., Carnero Rosell, A., Carrasco Kind, M., Carretero, J., Castander, F. J., Cawthon, R., Chang, C., Chen, N., Childfss, M., Choi, A., Conselice, C., Crittenden, R., Crocce, M., Cunha, C. E., D'Andrea, C. B., da Costa, L. N., Das, R., Davis, T. M., Davis, C., De Vicente, J., DePoy, D. L., DeRose, J., Desai, S., Diehl, H. T., Dietrich, J. P., Dodelson, S., Doel, P., Drlica-Wagner, A., Eifler, T. F., Elliott, A. E., Elsner, F., Elvin-Poole, J., Estrada, J., Evrard, A. E., Fang, Y., Fernandez, E., Ferté, A., Finley, D. A., Flaugher, B., Fosalba, P., Friedrich, O., Frieman, J., García-Bellido, J., Garcia-Fernandez, M., Gatti, M., Gaztanaga, E., Gerdes, D. W., Giannantonio, T., Gill, M. S. S., Glazebrook, K., Goldstein, D. A., Gruen, D., Gruendl, R. A., Gschwend, J., Gutierrez, G., Hamilton, S., Hartley, W. G., Hinton, S. R., Honscheid, K., Hoyle, B., Huterer, D., Jain, B., James, D. J., Jarvis, M., Jeltema, T., Johnson, M. D., Johnson, M. W. G., Kacprzak, T., Kent, S., Kim, A. G., King, A., Kirk, D., Kokron, N., Kovacs, A., Krause, E., Krawiec, C., Kremin, A., Kuehn, K., Kuhlmann, S., Kuropatkin, N., Lacasa, F., Lahav, O., Li, T. S., Liddle, A. R., Lidman, C., Lima, M., Lin, H., MacCrann, N., Maia, M. A. G., Makler, M., Manera, M., March, M., Marshall, J. L., Martini, P., McMahon, R. G., Melchior, P., Menanteau, F., Miquel, R., Miranda, V., Mudd, D., Muir, J., Möller, A., Neilsen, E., Nichol, R. C., Nord, B., Nugent, P., Ogando, R. L. C., Palmese, A., Peacock, J., Peiris, H. V., Peoples, J., Percival, W. J., Petravick, D., Plazas, A. A., Porredon, A., Prat, J., Pujol, A., Rau, M. M., Refregier, A., Ricker, P. M., Roe, N., Rollins, R. P., Romer, A. K., Roodman, A., Rosenfeld, R., Ross, A. J., Rozo, E., Rykoff, E. S., Sako, M., Salvador, A. I., Samuroff, S., Sánchez, C., Sanchez, E., Santiago, B., Scarpine, V., Schindler, R., Scolnic, D., Secco, L. F., Serrano, S., Sevilla-Noarbe, I., Sheldon, E., Smith, R. C., Smith, M., Smith, J., Soares-Santos, M., Sobreira, F., Suchyta, E., Tarle, G., Thomas, D., Troxel, M. A., Tucker, D. L., Tucker, B. E., Uddin, S. A., Varga, T. N., Vielzeuf, P., Vikram, V., Vivas, A. K., Walker, A. R., Wang, M., Wechsler, R. H., Weller, J., Wester, W., Wolf, R. C., Yanny, B., Yuan, F., Zenteno, A., Zhang, B., Zhang, Y., & Zuntz, J. (2017). Dark Energy Survey Year 1 Results: Cosmological Constraints from Galaxy Clustering and Weak Lensing. *ArXiv e-prints*.

Dickinson, C. (2014). BINGO - A novel method to detect BAOs using a total-power radio telescope. *ArXiv e-prints*.

D'Inverno, R. (1992). *Introducing Einstein's Relativity*. Clarendon Press, Oxford.

Douglas, J. N., Bash, F. N., Bozyan, F. A., Torrence, G. W., & Wolfe, C. (1996). The Texas Survey of Radio Sources Covering -35.5 degrees < declination < 71.5 degrees at 365 MHz. AJ, 111, 1945.

Eisenhauer, F., Perrin, G., Brandner, W., Straubmeier, C., Perraut, K., Amorim, A., Schöller, M., Gillessen, S., Kervella, P., Benisty, M., Araujo-Hauck, C., Jocou, L., Lima, J., Jakob, G., Haug, M., Clénet, Y., Henning, T., Eckart, A., Berger, J.-P., Garcia, P., Abuter, R., Kellner, S., Paumard, T., Hippler, S., Fischer, S., Moulin, T., Villate, J., Avila, G., Gräter, A., Lacour, S., Huber, A., Wiest, M., Nolot, A., Carvas, P., Dorn, R., Pfuhl, O., Gendron, E., Kendrew, S., Yazici, S., Anton, S., Jung, Y., Thiel, M., Choquet, É., Klein, R., Teixeira, P., Gitton, P., Moch, D., Vincent, F., Kudryavtseva, N., Ströbele, S., Sturm, S., Fédou, P., Lenzen, R., Jolley, P., Kister, C., Lapeyrère, V., Naranjo, V., Lucuix, C., Hofmann, R., Chapron, F., Neumann, U., Mehrgan, L., Hans, O., Rousset, G., Ramos, J., Suarez, M., Lederer, R., Reess, J.-M., Rohloff, R.-R., Haguenauer, P., Bartko, H., Sevin, A., Wagner, K., Lizon, J.-L., Rabien, S., Collin, C., Finger, G., Davies, R., Rouan, D., Wittkowski, M., Dodds-Eden, K., Ziegler, D., Cassaing, F., Bonnet, H., Casali, M., Genzel, R., & Lena, P. (2011). GRAVITY: Observing the Universe in Motion. *The Messenger*, 143, 16–24.

Eisenstein, D. J., Weinberg, D. H., Agol, E., Aihara, H., Allende Prieto, C., Anderson, S. F., Arns, J. A., Aubourg, É., Bailey, S., Balbinot, E., & et al. (2011). SDSS-III: Massive Spectroscopic Surveys of the Distant Universe, the Milky Way, and Extra-Solar Planetary Systems. AJ, 142, 72.

Fan, D., Budavári, T., Norris, R. P., & Hopkins, A. M. (2015). Matching radio catalogues with realistic geometry: application to SWIRE and ATLAS. MNRAS, 451, 1299–1305.

Fedorova, E., Sliusar, V. M., Zhdanov, V. I., Alexandrov, A. N., Del Popolo, A., & Surdej, J. (2016). Gravitational microlensing as a probe for dark matter clumps. MNRAS, 457, 4147–4159.

Fomalont, E. B., Kellermann, K. I., Cowie, L. L., & Barger, A. J. (2002). The Radio/Optical Morphology of Micro-Jansky Radio Sources. In *American Astronomical Society Meeting Abstracts*, volume 34 of *Bulletin of the American Astronomical Society* (pp. 1208).

Franceschini, A., Manners, J., Polletta, M. d. C., Lonsdale, C., Gonzalez-Solares, E., Surace, J., Shupe, D., Fang, F., Xu, C. K., Farrah, D., Berta, S., Rodighiero, G., Perez-Fournon, I., Hatziminaoglou, E., Smith, H. E., Siana, B., Rowan-Robinson, M., Nandra, K., Babbedge, T., Vaccari, M., Oliver, S., Wilkes, B., Owen, F., Padgett, D., Frayer, D., Jarrett, T., Masci, F., Stacey, G., Almaini, O., McMahon, R., Johnson, O., Lawrence, A., & Willott, C. (2005). A Complete Multiwavelength Characterization of Faint Chandra X-Ray Sources Seen in the Spitzer Wide-Area Infrared Extragalactic (SWIRE) Survey. AJ, 129, 2074–2101.

Garcia-Bellido, J. & Haugbølle, T. (2008). Confronting Lemaitre Tolman Bondi models with observational cosmology. J. Cosmology Astropart. Phys., 4, 003.

Garn, T., Green, D. A., Hales, S. E. G., Riley, J. M., & Alexander, P. (2007). Deep 610-MHz Giant Metrewave Radio Telescope observations of the Spitzer extragalactic First Look Survey field - I. Observations, data analysis and source catalogue. MNRAS, 376, 1251–1260.

Garn, T., Green, D. A., Riley, J. M., & Alexander, P. (2008). A 610-MHz survey of the ELAIS-N1 field with the Giant Metrewave Radio Telescope - observations, data analysis and source catalogue. MNRAS, 383, 75–85.

Garn, T., Green, D. A., Riley, J. M., & Alexander, P. (2009). The relationship between star formation rate and radio synchrotron luminosity at $0 < z < 2$. MNRAS, 397, 1101–1112.

Garnavich, P. M., Jha, S., Challis, P., Clocchiatti, A., Diercks, A., Filippenko, A. V., Gilliland, R. L., Hogan, C. J., Kirshner, R. P., Leibundgut, B., Phillips, M. M., Reiss, D., Riess, A. G., Schmidt, B. P., Schommer, R. A., Smith, R. C., Spyromilio, J., Stubbs, C., Suntzeff, N. B., Tonry, J., & Carroll, S. M. (1998). Supernova Limits on the Cosmic Equation of State. ApJ, 509, 74–79.

Garrett, M. A., Cordes, J. M., Deboer, D. R., Jonas, J. L., Rawlings, S., & Schilizzi, R. T. (2010). The Square Kilometre Array (SKA) - Phase 1 Design Concept. In *ISKAF2010 Science Meeting* (pp.1̃8).

Garsden, H., Girard, J. N., Starck, J. L., Corbel, S., Tasse, C., Woiselle, A., McKean, J. P., van Amesfoort, A. S., Anderson, J., Avruch, I. M., Beck, R., Bentum, M. J., Best, P., Breitling, F., Broderick, J., Brüggen, M., Butcher, H. R., Ciardi, B., de Gasperin, F., de Geus, E., de Vos, M., Duscha, S., Eislöffel, J., Engels, D., Falcke, H., Fallows, R. A., Fender, R., Ferrari, C., Frieswijk, W., Garrett, M. A., Grießmeier, J., Gunst, A. W., Hassall, T. E., Heald, G., Hoeft, M., Hörandel, J., van der Horst, A., Juette, E., Karastergiou, A., Kondratiev, V. I., Kramer, M., Kuniyoshi, M., Kuper, G., Mann, G., Markoff, S., McFadden, R., McKay-Bukowski, D., Mulcahy, D. D., Munk, H., Norden, M. J., Orru, E., Paas, H., Pandey-Pommier, M., Pandey, V. N., Pietka, G., Pizzo, R., Polatidis, A. G., Renting, A., Röttgering, H., Rowlinson, A., Schwarz, D., Sluman, J., Smirnov, O., Stappers, B. W., Steinmetz, M., Stewart, A., Swinbank, J., Tagger, M., Tang, Y., Tasse, C., Thoudam, S., Toribio, C., Vermeulen, R., Vocks, C., van Weeren, R. J., Wijnholds, S. J., Wise, M. W., Wucknitz, O., Yatawatta, S., Zarka, P., & Zensus, A. (2015). LOFAR sparse image reconstruction. A&A, 575, A90.

Gebhardt, K., Bender, R., Bower, G., Dressler, A., Faber, S. M., Filippenko, A. V., Green, R., Grillmair, C., Ho, L. C., Kormendy, J., Lauer, T. R., Magorrian, J., Pinkney, J., Richstone, D., & Tremaine, S. (2000). A Relationship between Nuclear Black Hole Mass and Galaxy Velocity Dispersion. ApJ, 539, L13–L16.

Gehrels, N. (1986). Confidence limits for small numbers of events in astrophysical data. ApJ, 303, 336–346.

Gregory, P. C. & Condon, J. J. (1991). The 87GB catalog of radio sources covering delta between O and + 75 deg at 4.85 GHz. ApJS, 75, 1011–1291.

Gunn, J. E., Siegmund, W. A., Mannery, E. J., Owen, R. E., Hull, C. L., Leger, R. F., Carey, L. N., Knapp, G. R., York, D. G., Boroski, W. N., Kent, S. M., Lupton, R. H., Rockosi, C. M., Evans, M. L., Waddell, P., Anderson, J. E., Annis, J., Barentine, J. C., Bartoszek, L. M., Bastian, S., Bracker, S. B., Brewington, H. J., Briegel, C. I., Brinkmann, J., Brown, Y. J., Carr, M. A., Czarapata, P. C., Drennan, C. C., Dombeck, T., Federwitz, G. R., Gillespie, B. A., Gonzales, C., Hansen, S. U., Harvanek, M., Hayes, J., Jordan, W., Kinney, E., Klaene, M., Kleinman, S. J., Kron, R. G., Kresinski, J., Lee, G., Limmongkol,

S., Lindenmeyer, C. W., Long, D. C., Loomis, C. L., McGehee, P. M., Mantsch, P. M., Neilsen, Jr., E. H., Neswold, R. M., Newman, P. R., Nitta, A., Peoples, Jr., J., Pier, J. R., Prieto, P. S., Prosapio, A., Rivetta, C., Schneider, D. P., Snedden, S., & Wang, S.-i. (2006). The 2.5 m Telescope of the Sloan Digital Sky Survey. AJ, 131, 2332–2359.

Guth, A. H. (1981). Inflationary universe: A possible solution to the horizon and flatness problems. Phys. Rev. D, 23, 347–356.

Hales, C. A., Murphy, T., Curran, J. R., Middelberg, E., Gaensler, B. M., & Norris, R. P. (2012). BLOBCAT: software to catalogue flood-filled blobs in radio images of total intensity and linear polarization. MNRAS, 425, 979–996.

Hales, S. E. G., Waldram, E. M., Rees, N., & Warner, P. J. (1995). A revised machine-readable source list for the Rees 38-MHz survey. MNRAS, 274, 447–451.

Harrison, I., Camera, S., Zuntz, J., & Brown, M. L. (2016). SKA weak lensing - I. Cosmological forecasts and the power of radio-optical cross-correlations. MNRAS, 463, 3674–3685.

Heymans, C. & Heavens, A. (2003). Weak gravitational lensing: reducing the contamination by intrinsic alignments. MNRAS, 339, 711–720.

Hildebrandt, H., Viola, M., Heymans, C., Joudaki, S., Kuijken, K., Blake, C., Erben, T., Joachimi, B., Klaes, D., Miller, L., Morrison, C. B., Nakajima, R., Verdoes Kleijn, G., Amon, A., Choi, A., Covone, G., de Jong, J. T. A., Dvornik, A., Fenech Conti, I., Grado, A., Harnois-Déraps, J., Herbonnet, R., Hoekstra, H., Köhlinger, F., McFarland, J., Mead, A., Merten, J., Napolitano, N., Peacock, J. A., Radovich, M., Schneider, P., Simon, P., Valentijn, E. A., van den Busch, J. L., van Uitert, E., & Van Waerbeke, L. (2017). KiDS-450: cosmological parameter constraints from tomographic weak gravitational lensing. MNRAS, 465, 1454–1498.

Hine, R. G. & Longair, M. S. (1979). Optical spectra of 3CR radio galaxies. MNRAS, 188, 111–130.

Hirabayashi, H., Hirosawa, H., Kobayashi, H., Murata, Y., Edwards, P. G., Fomalont, E. B., Fujisawa, K., Ichikawa, T., Kii, T., Lovell, J. E. J., Moellenbrock, G. A., Okayasu, R., Inoue, M., Kawaguchi, N., Kameno, S., Shibata, K. M., Asaki, Y., Bushimata, T., Enome, S., Horiuchi, S., Miyaji, T., Umemoto, T., Migenes, V., Wajima, K., Nakajima, J., Morimoto, M., Ellis, J., Meier, D. L., Murphy, D. W., Preston, R. A., Smith, J. G., Tingay, S. J., Traub, D. L., Wietfeldt, R. D., Benson, J. M., Claussen, M. J., Flatters, C., Romney, J. D., Ulvestad, J. S., D'Addario, L. R., Langston, G. I., Minter, A. H., Carlson, B. R., Dewdney, P. E., Jauncey, D. L., Reynolds, J. E., Taylor, A. R., McCulloch, P. M., Cannon, W. H., Gurvits, L. I., Mioduszewski, A. J., Schilizzi, R. T., & Booth, R. S. (1998). Overview and Initial Results of the Very Long Baseline Interferometry Space Observatory Programme. *Science*, 281, 1825.

Högbom, J. A. (1974). Aperture Synthesis with a Non-Regular Distribution of Interferometer Baselines. A&AS, 15, 417.

Holtzman, J. A. (1989). Microwave background anisotropies and large-scale structure in universes with cold dark matter, baryons, radiation, and massive and massless neutrinos. ApJS, 71, 1–24.

Hopkins, A. M., Afonso, J., Chan, B., Cram, L. E., Georgakakis, A., & Mobasher, B. (2003). The Phoenix Deep Survey: The 1.4 GHz Microjansky Catalog. AJ, 125, 465–477.

Howell, D. A., Sullivan, M., Nugent, P. E., Ellis, R. S., Conley, A. J., Le Borgne, D., Carlberg, R. G., Guy, J., Balam, D., Basa, S., Fouchez, D., Hook, I. M., Hsiao, E. Y., Neill, J. D., Pain, R., Perrett, K. M., & Pritchet, C. J. (2006). The type Ia supernova SNLS-03D3bb from a super-Chandrasekhar-mass white dwarf star. Nature, 443, 308–311.

Hubble, E. (1929). A Relation between Distance and Radial Velocity among Extra-Galactic Nebulae. *Proceedings of the National Academy of Science*, 15, 168–173.

Intema, H. T., van der Tol, S., Cotton, W. D., Cohen, A. S., van Bemmel, I. M., & Röttgering, H. J. A. (2009). Ionospheric calibration of low frequency radio interferometric observations using the peeling scheme. I. Method description and first results. A&A, 501, 1185–1205.

Ivezić, Ž., Menou, K., Knapp, G. R., Strauss, M. A., Lupton, R. H., Vanden Berk, D. E., Richards, G. T., Tremonti, C., Weinstein, M. A., Anderson, S., Bahcall, N. A., Becker, R. H., Bernardi, M., Blanton, M., Eisenstein, D., Fan, X., Finkbeiner, D., Finlator, K., Frieman, J., Gunn, J. E., Hall, P. B., Kim, R. S. J., Kinkhabwala, A., Narayanan, V. K., Rockosi, C. M., Schlegel, D., Schneider, D. P., Strateva, I., SubbaRao, M., Thakar, A. R., Voges, W., White, R. L., Yanny, B., Brinkmann, J., Doi, M., Fukugita, M., Hennessy, G. S., Munn, J. A., Nichol, R. C., & York, D. G. (2002). Optical and Radio Properties of Extragalactic Sources Observed by the FIRST Survey and the Sloan Digital Sky Survey. AJ, 124, 2364–2400.

Jacobs, C., Glazebrook, K., Collett, T., More, A., & McCarthy, C. (2017). Finding strong lenses in CFHTLS using convolutional neural networks. *ArXiv e-prints*.

Jansky, K. G. (1935). A note on the source of interstellar interference. *Proc. IRE*, 23, 1158.

Jarvis, M. J. (2007). Science with the Next Generation of Radio Surveys from LOFAR to the SKA. In J. Afonso, H. C. Ferguson, B. Mobasher, & R. Norris (Eds.), *Deepest Astronomical Surveys*, volume 380 of *Astronomical Society of the Pacific Conference Series* (pp. 251).

Joachimi, B. & Schneider, P. (2008). The removal of shear-ellipticity correlations from the cosmic shear signal via nulling techniques. A&A, 488, 829–843.

Jones, T. A., Swinbank, A. M., Ellis, R. S., Richard, J., & Stark, D. P. (2010). Resolved spectroscopy of gravitationally lensed galaxies: recovering coherent velocity fields in subluminous z ˜ 2-3 galaxies. MNRAS, 404, 1247–1262.

Kaiser, N. (1992). Weak gravitational lensing of distant galaxies. ApJ, 388, 272–286.

Kaiser, N., Squires, G., & Broadhurst, T. (1995). A Method for Weak Lensing Observations. ApJ, 449, 460.

Kauffmann, G., Heckman, T. M., Tremonti, C., Brinchmann, J., Charlot, S., White, S. D. M., Ridgway, S. E., Brinkmann, J., Fukugita, M., Hall, P. B., Ivezić, Ž., Richards, G. T., & Schneider, D. P. (2003). The host galaxies of active galactic nuclei. MNRAS, 346, 1055–1077.

Kazemi, S., Yatawatta, S., Zaroubi, S., Lampropoulos, P., de Bruyn, A. G., Koopmans, L. V. E., & Noordam, J. (2011). Radio interferometric calibration using the SAGE algorithm. MNRAS, 414, 1656–1666.

Kilbinger, M. (2015). Cosmology with cosmic shear observations: a review. *Reports on Progress in Physics*, 78, 086901.

Kim, J.-W., Edge, A. C., Wake, D. A., Gonzalez-Perez, V., Baugh, C. M., Lacey, C. G., Yamada, T., Sato, Y., Burgett, W. S., Chambers, K. C., Price, P. A., Foucaud, S., Draper, P., & Kaiser, N. (2014). Clustering of extremely red objects in Elais-N1 from the UKIDSS DXS with optical photometry from Pan-STARRS 1 and Subaru. MNRAS, 438, 825–840.

Kitaura, F.-S., Rodríguez-Torres, S., Chuang, C.-H., Zhao, C., Prada, F., Gil-Marín, H., Guo, H., Yepes, G., Klypin, A., Scóccola, C. G., Tinker, J., McBride, C., Reid, B., Sánchez, A. G., Salazar-Albornoz, S., Grieb, J. N., Vargas-Magana, M., Cuesta, A. J., Neyrinck, M., Beutler, F., Comparat, J., Percival, W. J., & Ross, A. (2016). The clustering of galaxies in the SDSS-III Baryon Oscillation Spectroscopic Survey: mock galaxy catalogues for the BOSS Final Data Release. MNRAS, 456, 4156–4173.

Kitaura, F.-S., Yepes, G., & Prada, F. (2014). Modelling baryon acoustic oscillations with perturbation theory and stochastic halo biasing. MNRAS, 439, L21–L25.

Kitching, T. D., Balan, S. T., Bridle, S., Cantale, N., Courbin, F., Eifler, T., Gentile, M., Gill, M. S. S., Harmeling, S., Heymans, C., Hirsch, M., Honscheid, K., Kacprzak, T., Kirkby, D., Margala, D., Massey, R. J., Melchior, P., Nurbaeva, G., Patton, K., Rhodes, J., Rowe, B. T. P., Taylor, A. N., Tewes, M., Viola, M., Witherick, D., Voigt, L., Young, J., & Zuntz, J. (2012). Image analysis for cosmology: results from the GREAT10 Galaxy Challenge. MNRAS, 423, 3163–3208.

Kitching, T. D., Heavens, A. F., Alsing, J., Erben, T., Heymans, C., Hildebrandt, H., Hoekstra, H., Jaffe, A., Kiessling, A., Mellier, Y., Miller, L., van Waerbeke, L., Benjamin, J., Coupon, J., Fu, L., Hudson, M. J., Kilbinger, M., Kuijken, K., Rowe, B. T. P., Schrabback, T., Semboloni, E., & Velander, M. (2014). 3D cosmic shear: cosmology from CFHTLenS. MNRAS, 442, 1326–1349.

Komatsu, E., Smith, K. M., Dunkley, J., Bennett, C. L., Gold, B., Hinshaw, G., Jarosik, N., Larson, D., Nolta, M. R., Page, L., Spergel, D. N., Halpern, M., Hill, R. S., Kogut, A., Limon, M., Meyer, S. S., Odegard, N., Tucker, G. S., Weiland, J. L., Wollack, E., & Wright, E. L. (2011). Seven-year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Cosmological Interpretation. ApJS, 192, 18.

Larson, D., Dunkley, J., Hinshaw, G., Komatsu, E., Nolta, M. R., Bennett, C. L., Gold, B., Halpern, M., Hill, R. S., Jarosik, N., Kogut, A., Limon, M., Meyer, S. S., Odegard, N., Page, L., Smith, K. M., Spergel, D. N., Tucker, G. S., Weiland, J. L., Wollack, E., & Wright, E. L. (2011). Seven-year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Power Spectra and WMAP-derived Parameters. ApJS, 192, 16.

Lawrence, A., Warren, S. J., Almaini, O., Edge, A. C., Hambly, N. C., Jameson, R. F., Lucas, P., Casali, M., Adamson, A., Dye, S., Emerson, J. P., Foucaud, S., Hewett, P., Hirst, P., Hodgkin, S. T., Irwin, M. J., Lodieu, N., McMahon, R. G., Simpson, C., Smail, I.,

Mortlock, D., & Folger, M. (2007). The UKIRT Infrared Deep Sky Survey (UKIDSS). MNRAS, 379, 1599–1617.

Leavitt, H. S. (1908). 1777 variables in the Magellanic Clouds. *Annals of Harvard College Observatory*, 60, 87–108.3.

Liddle, A. (2015). *An Introduction to Modern Cosmology*. Wiley, third edition.

Lonsdale, C. J. (2007). Science with the MWA. In *From Planets to Dark Energy: the Modern Radio Universe* (pp. 146).

Lonsdale, C. J., Smith, H. E., Rowan-Robinson, M., Surace, J., Shupe, D., Xu, C., Oliver, S., Padgett, D., Fang, F., Conrow, T., Franceschini, A., Gautier, N., Griffin, M., Hacking, P., Masci, F., Morrison, G., O'Linger, J., Owen, F., Pérez-Fournon, I., Pierre, M., Puetter, R., Stacey, G., Castro, S., Polletta, M. d. C., Farrah, D., Jarrett, T., Frayer, D., Siana, B., Babbedge, T., Dye, S., Fox, M., Gonzalez-Solares, E., Salaman, M., Berta, S., Condon, J. J., Dole, H., & Serjeant, S. (2003). SWIRE: The SIRTF Wide-Area Infrared Extragalactic Survey. PASP, 115, 897–927.

Maggiore, M. (2000). Gravitational wave experiments and early universe cosmology. Phys. Rep., 331, 283–367.

Mangano, G., Miele, G., Pastor, S., & Peloso, M. (2002). A precision calculation of the effective number of cosmological neutrinos. *Physics Letters B*, 534, 8–16.

Manners, J. C., Johnson, O., Almaini, O., Willott, C. J., Gonzalez-Solares, E., Lawrence, A., Mann, R. G., Perez-Fournon, I., Dunlop, J. S., McMahon, R. G., Oliver, S. J., Rowan-Robinson, M., & Serjeant, S. (2003). The ELAIS deep X-ray survey - I. Chandra source catalogue and first results. MNRAS, 343, 293–305.

Maraston, C. (2005). Evolutionary population synthesis: models, analysis of the ingredients and application to high-z galaxies. MNRAS, 362, 799–825.

Martí-Vidal, I., Vlemmings, W. H. T., Muller, S., & Casey, S. (2014). UVMULTIFIT: A versatile tool for fitting astronomical radio interferometric data. A&A, 563, A136.

Mauduit, J.-C., Lacy, M., Farrah, D., Surace, J. A., Jarvis, M., Oliver, S., Maraston, C., Vaccari, M., Marchetti, L., Zeimann, G., Gonzáles-Solares, E. A., Pforr, J., Petric, A. O., Henriques, B., Thomas, P. A., Afonso, J., Rettura, A., Wilson, G., Falder, J. T., Geach, J. E., Huynh, M., Norris, R. P., Seymour, N., Richards, G. T., Stanford, S. A., Alexander, D. M., Becker, R. H., Best, P. N., Bizzocchi, L., Bonfield, D., Castro, N., Cava, A., Chapman, S., Christopher, N., Clements, D. L., Covone, G., Dubois, N., Dunlop, J. S., Dyke, E., Edge, A., Ferguson, H. C., Foucaud, S., Franceschini, A., Gal, R. R., Grant, J. K., Grossi, M., Hatziminaoglou, E., Hickey, S., Hodge, J. A., Huang, J.-S., Ivison, R. J., Kim, M., LeFevre, O., Lehnert, M., Lonsdale, C. J., Lubin, L. M., McLure, R. J., Messias, H., Martínez-Sansigre, A., Mortier, A. M. J., Nielsen, D. M., Ouchi, M., Parish, G., Perez-Fournon, I., Pierre, M., Rawlings, S., Readhead, A., Ridgway, S. E., Rigopoulou, D., Romer, A. K., Rosebloom, I. G., Rottgering, H. J. A., Rowan-Robinson, M., Sajina, A., Simpson, C. J., Smail, I., Squires, G. K., Stevens, J. A., Taylor, R., Trichas, M., Urrutia, T., van Kampen, E., Verma, A., & Xu, C. K. (2012). The Spitzer Extragalactic Representative Volume Survey (SERVS): Survey Definition and Goals. PASP, 124, 714.

McAlpine, K., Smith, D. J. B., Jarvis, M. J., Bonfield, D. G., & Fleuren, S. (2012). The likelihood ratio as a tool for radio continuum surveys with Square Kilometre Array precursor telescopes. MNRAS, 423, 132–140.

McMahon, R. G., White, R. L., Helfand, D. J., & Becker, R. H. (2002). Optical Counterparts for 70,000 Radio Sources: APM Identifications for the FIRST Radio Survey. *The Astrophysical Journal Supplement Series*, 143, 1–23.

McMullin, J. P., Waters, B., Schiebel, D., Young, W., & Golap, K. (2007). CASA Architecture and Applications. In R. A. Shaw, F. Hill, & D. J. Bell (Eds.), *Astronomical Data Analysis Software and Systems XVI*, volume 376 of *Astronomical Society of the Pacific Conference Series* (pp. 127).

Menzel, M.-L., Merloni, A., Georgakakis, A., Salvato, M., Aubourg, E., Brandt, W. N., Brusa, M., Buchner, J., Dwelly, T., Nandra, K., Pâris, I., Petitjean, P., & Schwope, A. (2016). A spectroscopic survey of X-ray-selected AGNs in the northern XMM-XXL field. MNRAS, 457, 110–132.

Miller, L., Heymans, C., Kitching, T. D., van Waerbeke, L., Erben, T., Hildebrandt, H., Hoekstra, H., Mellier, Y., Rowe, B. T. P., Coupon, J., Dietrich, J. P., Fu, L., Harnois-Déraps, J., Hudson, M. J., Kilbinger, M., Kuijken, K., Schrabback, T., Semboloni, E., Vafaei, S., & Velander, M. (2013). Bayesian galaxy shape measurement for weak lensing surveys - III. Application to the Canada-France-Hawaii Telescope Lensing Survey. MNRAS, 429, 2858–2880.

Miyazaki, S., Komiyama, Y., Nakaya, H., Kamata, Y., Doi, Y., Hamana, T., Karoji, H., Furusawa, H., Kawanomoto, S., Morokuma, T., Ishizuka, Y., Nariai, K., Tanaka, Y., Uraguchi, F., Utsumi, Y., Obuchi, Y., Okura, Y., Oguri, M., Takata, T., Tomono, D., Kurakami, T., Namikawa, K., Usuda, T., Yamanoi, H., Terai, T., Uekiyo, H., Yamada, Y., Koike, M., Aihara, H., Fujimori, Y., Mineo, S., Miyatake, H., Yasuda, N., Nishizawa, J., Saito, T., Tanaka, M., Uchida, T., Katayama, N., Wang, S.-Y., Chen, H.-Y., Lupton, R., Loomis, C., Bickerton, S., Price, P., Gunn, J., Suzuki, H., Miyazaki, Y., Muramatsu, M., Yamamoto, K., Endo, M., Ezaki, Y., Itoh, N., Miwa, Y., Yokota, H., Matsuda, T., Ebinuma, R., & Takeshi, K. (2012). Hyper Suprime-Cam. In *Ground-based and Airborne Instrumentation for Astronomy IV*, volume 8446 of Proc. SPIE (pp. 84460Z).

Mohan, N. & Rafferty, D. (2015). PyBDSM: Python Blob Detection and Source Measurement. Astrophysics Source Code Library.

Morales, M. F. (2006). A Technique for Weak Lensing with Velocity Maps: Eliminating Ellipticity Noise in H I Radio Observations. ApJ, 650, L21–L23.

Nisbet, D. (2016). Automated cross-id with multi-band likelihoods. https://www.dropbox.com/s/s0qq1qw97oefy89/DavidNisbet_WEAVE_LOFAR.pdf?dl=0.

Noordam, J. E. (2004). LOFAR calibration challenges. In J. M. Oschmann, Jr. (Ed.), *Ground-based Telescopes*, volume 5489 of Proc. SPIE (pp. 817–825).

Norris, R. P., Afonso, J., Appleton, P. N., Boyle, B. J., Ciliegi, P., Croom, S. M., Huynh, M. T., Jackson, C. A., Koekemoer, A. M., Lonsdale, C. J., Middelberg, E., Mobasher, B., Oliver, S. J., Polletta, M., Siana, B. D., Smail, I., & Voronkov, M. A. (2006). Deep

ATLAS Radio Observations of the Chandra Deep Field-South/Spitzer Wide-Area Infrared Extragalactic Field. AJ, 132, 2409–2423.

Offringa, A. R., van de Gronde, J. J., & Roerdink, J. B. T. M. (2012). A morphological algorithm for improving radio-frequency interference detection. A&A, 539, A95.

Oliver, S. J., Bock, J., Altieri, B., Amblard, A., Arumugam, V., Aussel, H., Babbedge, T., Beelen, A., Béthermin, M., Blain, A., Boselli, A., Bridge, C., Brisbin, D., Buat, V., Burgarella, D., Castro-Rodríguez, N., Cava, A., Chanial, P., Cirasuolo, M., Clements, D. L., Conley, A., Conversi, L., Cooray, A., Dowell, C. D., Dubois, E. N., Dwek, E., Dye, S., Eales, S., Elbaz, D., Farrah, D., Feltre, A., Ferrero, P., Fiolet, N., Fox, M., Franceschini, A., Gear, W., Giovannoli, E., Glenn, J., Gong, Y., González Solares, E. A., Griffin, M., Halpern, M., Harwit, M., Hatziminaoglou, E., Heinis, S., Hurley, P., Hwang, H. S., Hyde, A., Ibar, E., Ilbert, O., Isaak, K., Ivison, R. J., Lagache, G., Le Floc'h, E., Levenson, L., Faro, B. L., Lu, N., Madden, S., Maffei, B., Magdis, G., Mainetti, G., Marchetti, L., Marsden, G., Marshall, J., Mortier, A. M. J., Nguyen, H. T., O'Halloran, B., Omont, A., Page, M. J., Panuzzo, P., Papageorgiou, A., Patel, H., Pearson, C. P., Pérez-Fournon, I., Pohlen, M., Rawlings, J. I., Raymond, G., Rigopoulou, D., Riguccini, L., Rizzo, D., Rodighiero, G., Roseboom, I. G., Rowan-Robinson, M., Sánchez Portal, M., Schulz, B., Scott, D., Seymour, N., Shupe, D. L., Smith, A. J., Stevens, J. A., Symeonidis, M., Trichas, M., Tugwell, K. E., Vaccari, M., Valtchanov, I., Vieira, J. D., Viero, M., Vigroux, L., Wang, L., Ward, R., Wardlow, J., Wright, G., Xu, C. K., & Zemcov, M. (2012). The Herschel Multi-tiered Extragalactic Survey: HerMES. MNRAS, 424, 1614–1635.

Padovani, P., Mainieri, V., Tozzi, P., Kellermann, K. I., Fomalont, E. B., Miller, N., Rosati, P., & Shaver, P. (2007). The micro-Jy Radio Source Population: the VLA-CDFS View. In J. Afonso, H. C. Ferguson, B. Mobasher, & R. Norris (Eds.), *Deepest Astronomical Surveys*, volume 380 of *Astronomical Society of the Pacific Conference Series* (pp. 205).

Pandey, V. N., van Zwieten, J. E., de Bruyn, A. G., & Nijboer, R. (2009). Calibrating LOFAR using the Black Board Selfcal System. In D. J. Saikia, D. A. Green, Y. Gupta, & T. Venturi (Eds.), *The Low-Frequency Radio Universe*, volume 407 of *Astronomical Society of the Pacific Conference Series* (pp. 384).

Pâris, I., Petitjean, P., Aubourg, É., Ross, N. P., Myers, A. D., Streblyanska, A., Bailey, S., Hall, P. B., Strauss, M. A., Anderson, S. F., Bizyaev, D., Borde, A., Brinkmann, J., Bovy, J., Brandt, W. N., Brewington, H., Brownstein, J. R., Cook, B. A., Ebelke, G., Fan, X., Filiz Ak, N., Finley, H., Font-Ribera, A., Ge, J., Hamann, F., Ho, S., Jiang, L., Kinemuchi, K., Malanushenko, E., Malanushenko, V., Marchante, M., McGreer, I. D., McMahon, R. G., Miralda-Escudé, J., Muna, D., Noterdaeme, P., Oravetz, D., Palanque-Delabrouille, N., Pan, K., Perez-Fournon, I., Pieri, M., Riffel, R., Schlegel, D. J., Schneider, D. P., Simmons, A., Viel, M., Weaver, B. A., Wood-Vasey, W. M., Yèche, C., & York, D. G. (2014). The Sloan Digital Sky Survey quasar catalog: tenth data release. A&A, 563, A54.

Pâris, I., Petitjean, P., Ross, N. P., Myers, A. D., Aubourg, É., Streblyanska, A., Bailey, S., Armengaud, É., Palanque-Delabrouille, N., Yèche, C., Hamann, F., Strauss, M. A., Albareti, F. D., Bovy, J., Bizyaev, D., Niel Brandt, W., Brusa, M., Buchner, J., Comparat, J., Croft, R. A. C., Dwelly, T., Fan, X., Font-Ribera, A., Ge, J., Georgakakis, A., Hall, P. B., Jiang, L., Kinemuchi, K., Malanushenko, E., Malanushenko, V., McMahon, R. G., Menzel, M.-L., Merloni, A., Nandra, K., Noterdaeme, P., Oravetz, D., Pan, K., Pieri, M. M., Prada,

F., Salvato, M., Schlegel, D. J., Schneider, D. P., Simmons, A., Viel, M., Weinberg, D. H., & Zhu, L. (2017). The Sloan Digital Sky Survey Quasar Catalog: Twelfth data release. A&A, 597, A79.

Pascual, S., Gallego, J., Aragón-Salamanca, A., & Zamorano, J. (2001). Hα emitting galaxies and the star formation rate density at z =˜ 0.24. A&A, 379, 798–806.

Patel, P., Abdalla, F. B., Bacon, D. J., Rowe, B., Smirnov, O. M., & Beswick, R. J. (2014). Weak lensing measurements in simulations of radio images. MNRAS, 444, 2893–2909.

Patel, P., Bacon, D. J., Beswick, R. J., Muxlow, T. W. B., & Hoyle, B. (2010). Radio weak gravitational lensing with VLA and MERLIN. MNRAS, 401, 2572–2586.

Pearson, T. J. & C-BASS Collaboration (2016). C-BASS: The C-Band All Sky Survey. In *American Astronomical Society Meeting Abstracts*, volume 228 of *American Astronomical Society Meeting Abstracts* (pp. 301.04).

Peebles, P. J. E. & Yu, J. T. (1970). Primeval Adiabatic Perturbation in an Expanding Universe. ApJ, 162, 815.

Penzias, A. A. & Wilson, R. W. (1965). A Measurement of Excess Antenna Temperature at 4080 Mc/s. ApJ, 142, 419–421.

Perlmutter, S., Aldering, G., della Valle, M., Deustua, S., Ellis, R. S., Fabbro, S., Fruchter, A., Goldhaber, G., Groom, D. E., Hook, I. M., Kim, A. G., Kim, M. Y., Knop, R. A., Lidman, C., McMahon, R. G., Nugent, P., Pain, R., Panagia, N., Pennypacker, C. R., Ruiz-Lapuente, P., Schaefer, B., & Walton, N. (1998). Discovery of a supernova explosion at half the age of the universe. Nature, 391, 51.

Phillips, M. M. (1993). The absolute magnitudes of Type IA supernovae. ApJ, 413, L105–L108.

Planck Collaboration, Ade, P. A. R., Aghanim, N., Armitage-Caplan, C., Arnaud, M., Ashdown, M., Atrio-Barandela, F., Aumont, J., Baccigalupi, C., Banday, A. J., & et al. (2014). Planck 2013 results. XVI. Cosmological parameters. A&A, 571, A16.

Planck Collaboration, Ade, P. A. R., Aghanim, N., Arnaud, M., Ashdown, M., Aumont, J., Baccigalupi, C., Banday, A. J., Barreiro, R. B., Bartlett, J. G., & et al. (2016). Planck 2015 results. XIII. Cosmological parameters. A&A, 594, A13.

Prandoni, I., Gregorini, L., Parma, P., de Ruiter, H. R., Vettolani, G., Wieringa, M. H., & Ekers, R. D. (2001). The ATESP radio survey. III. Source counts. A&A, 365, 392–399.

Pritchard, J., Ichiki, K., Mesinger, A., Metcalf, R. B., Pourtsidou, A., Santos, M., Abdalla, F. B., Chang, T. C., Chen, X., Weller, J., & Zaroubi, S. (2015). Cosmology from EoR/Cosmic Dawn with the SKA. *Advancing Astrophysics with the Square Kilometre Array (AASKA14)*, (pp.1̃2).

Refregier, A. (2003). Shapelets - I. A method for image analysis. MNRAS, 338, 35–47.

Refregier, A. & Bacon, D. (2003). Shapelets - II. A method for weak lensing measurements. MNRAS, 338, 48–56.

Rengelink, R. B., Tang, Y., de Bruyn, A. G., Miley, G. K., Bremer, M. N., Roettgering, H. J. A., & Bremer, M. A. R. (1997). The Westerbork Northern Sky Survey (WENSS), I. A 570 square degree Mini-Survey around the North Ecliptic Pole. A&AS, 124.

Riess, A. G., Filippenko, A. V., Challis, P., Clocchiatti, A., Diercks, A., Garnavich, P. M., Gilliland, R. L., Hogan, C. J., Jha, S., Kirshner, R. P., Leibundgut, B., Phillips, M. M., Reiss, D., Schmidt, B. P., Schommer, R. A., Smith, R. C., Spyromilio, J., Stubbs, C., Suntzeff, N. B., & Tonry, J. (1998). Observational Evidence from Supernovae for an Accelerating Universe and a Cosmological Constant. AJ, 116, 1009–1038.

Riess, A. G., Macri, L. M., Hoffmann, S. L., Scolnic, D., Casertano, S., Filippenko, A. V., Tucker, B. E., Reid, M. J., Jones, D. O., Silverman, J. M., Chornock, R., Challis, P., Yuan, W., Brown, P. J., & Foley, R. J. (2016). A 2.4% Determination of the Local Value of the Hubble Constant. ApJ, 826, 56.

Riess, A. G., Press, W. H., & Kirshner, R. P. (1996). A Precise Distance Indicator: Type IA Supernova Multicolor Light-Curve Shapes. ApJ, 473, 88.

Rivi, M., Miller, L., Makhathini, S., & Abdalla, F. B. (2016a). Radio weak lensing shear measurement in the visibility domain - I. Methodology. MNRAS, 463, 1881–1890.

Rivi, M., Miller, L., Makhathini, S., & Batoni Abdalla, F. (2016b). RadioLensfit: Bayesian weak lensing measurement in the visibility domain. *ArXiv e-prints*.

Rohlfs, K. & Wilson, T. L. (2000). *Tools of radio astronomy*.

Ross, A. J., Beutler, F., Chuang, C.-H., Pellejero-Ibanez, M., Seo, H.-J., Vargas-Magaña, M., Cuesta, A. J., Percival, W. J., Burden, A., Sánchez, A. G., Grieb, J. N., Reid, B., Brownstein, J. R., Dawson, K. S., Eisenstein, D. J., Ho, S., Kitaura, F.-S., Nichol, R. C., Olmstead, M. D., Prada, F., Rodríguez-Torres, S. A., Saito, S., Salazar-Albornoz, S., Schneider, D. P., Thomas, D., Tinker, J., Tojeiro, R., Wang, Y., White, M., & Zhao, G.-b. (2017). The clustering of galaxies in the completed SDSS-III Baryon Oscillation Spectroscopic Survey: observational systematics and baryon acoustic oscillations in the correlation function. MNRAS, 464, 1168–1191.

Ross, A. J., Percival, W. J., Sánchez, A. G., Samushia, L., Ho, S., Kazin, E., Manera, M., Reid, B., White, M., Tojeiro, R., McBride, C. K., Xu, X., Wake, D. A., Strauss, M. A., Montesano, F., Swanson, M. E. C., Bailey, S., Bolton, A. S., Dorta, A. M., Eisenstein, D. J., Guo, H., Hamilton, J.-C., Nichol, R. C., Padmanabhan, N., Prada, F., Schlegel, D. J., Magaña, M. V., Zehavi, I., Blanton, M., Bizyaev, D., Brewington, H., Cuesta, A. J., Malanushenko, E., Malanushenko, V., Oravetz, D., Parejko, J., Pan, K., Schneider, D. P., Shelden, A., Simmons, A., Snedden, S., & Zhao, G.-b. (2012). The clustering of galaxies in the SDSS-III Baryon Oscillation Spectroscopic Survey: analysis of potential systematics. MNRAS, 424, 564–590.

Rottgering, H. J. A. (2010). LOFAR and the low frequency Universe. In *ISKAF2010 Science Meeting* (pp.5̃0).

Rubin, V. C. & Ford, Jr., W. K. (1970). Rotation of the Andromeda Nebula from a Spectroscopic Survey of Emission Regions. ApJ, 159, 379.

Sachs, R. K. & Wolfe, A. M. (1967). Perturbations of a Cosmological Model and Angular Variations of the Microwave Background. ApJ, 147, 73.

Saripalli, L. (2012). Understanding the Fanaroff-Riley Radio Galaxy Classification. AJ, 144, 85.

Schmidt, B. P., Suntzeff, N. B., Phillips, M. M., Schommer, R. A., Clocchiatti, A., Kirshner, R. P., Garnavich, P., Challis, P., Leibundgut, B., Spyromilio, J., Riess, A. G., Filippenko, A. V., Hamuy, M., Smith, R. C., Hogan, C., Stubbs, C., Diercks, A., Reiss, D., Gilliland, R., Tonry, J., Maza, J., Dressler, A., Walsh, J., & Ciardullo, R. (1998). The High-Z Supernova Search: Measuring Cosmic Deceleration and Global Curvature of the Universe Using Type IA Supernovae. ApJ, 507, 46–63.

Schwarz, U. J. (1978). Mathematical-statistical Description of the Iterative Beam Removing Technique (Method CLEAN). A&A, 65, 345.

Seitz, C. & Schneider, P. (1995). Steps towards nonlinear cluster inversion through gravitational distortions II. Generalization of the Kaiser and Squires method. A&A, 297, 287.

Seymour, N., McHardy, I. M., & Gunn, K. F. (2004). Radio observations of the 13hxmm–newton/rosat deep x-ray survey area. *Monthly Notices of the Royal Astronomical Society*, 352(1), 131–141.

Sheldon, E. (2015). NGMIX: Gaussian mixture models for 2D images. Astrophysics Source Code Library.

Silk, J. I. (1968). *The Formation of Galaxies.* PhD thesis, HARVARD UNIVERSITY.

Simpson, C., Martínez-Sansigre, A., Rawlings, S., Ivison, R., Akiyama, M., Sekiguchi, K., Takata, T., Ueda, Y., & Watson, M. (2006). Radio imaging of the subaru/xmm–newton deep field – i. the 100-μjy catalogue, optical identifications, and the nature of the faint radio source population. *Monthly Notices of the Royal Astronomical Society*, 372(2), 741–757.

Sirothia, S. K., Dennefeld, M., Saikia, D. J., Dole, H., Ricquebourg, F., & Roland, J. (2009). 325-MHz observations of the ELAIS-N1 field using the Giant Metrewave Radio Telescope. MNRAS, 395, 269–281.

Smith, D. J. B., Best, P. N., Duncan, K. J., Hatch, N. A., Jarvis, M. J., Röttgering, H. J. A., Simpson, C. J., Stott, J. P., Cochrane, R. K., Coppin, K. E., Dannerbauer, H., Davis, T. A., Geach, J. E., Hale, C. L., Hardcastle, M. J., Hatfield, P. W., Houghton, R. C. W., Maddox, N., McGee, S. L., Morabito, L., Nisbet, D., Pandey-Pommier, M., Prandoni, I., Saxena, A., Shimwell, T. W., Tarr, M., van Bemmel, I., Verma, A., White, G. J., & Williams, W. L. (2016). The WEAVE-LOFAR Survey. *ArXiv e-prints*.

Smoot, G. F., Bennett, C. L., Kogut, A., Aymon, J., Backus, C., de Amici, G., Galuk, K., Jackson, P. D., Keegstra, P., Rokke, L., Tenorio, L., Torres, S., Gulkis, S., Hauser, M. G., Janssen, M. A., Mather, J. C., Weiss, R., Wilkinson, D. T., Wright, E. L., Boggess, N. W., Cheng, E. S., Kelsall, T., Lubin, P., Meyer, S., Moseley, S. H., Murdock, T. L., Shafer, R. A., & Silverberg, R. F. (1991). Preliminary results from the COBE differential

microwave radiometers - Large angular scale isotropy of the cosmic microwave background. ApJ, 371, L1–L5.

Sunyaev, R. A. & Zeldovich, Y. B. (1970). Small-Scale Fluctuations of Relic Radiation. Ap&SS, 7, 3–19.

Sutherland, W. & Saunders, W. (1992). On the likelihood ratio for source identification. MNRAS, 259, 413–420.

Suyu, S. H., Auger, M. W., Hilbert, S., Marshall, P. J., Tewes, M., Treu, T., Fassnacht, C. D., Koopmans, L. V. E., Sluse, D., Blandford, R. D., Courbin, F., & Meylan, G. (2013). Two Accurate Time-delay Distances from Strong Lensing: Implications for Cosmology. ApJ, 766, 70.

Suzuki, N., Rubin, D., Lidman, C., Aldering, G., Amanullah, R., Barbary, K., Barrientos, L. F., Botyanszki, J., Brodwin, M., Connolly, N., Dawson, K. S., Dey, A., Doi, M., Donahue, M., Deustua, S., Eisenhardt, P., Ellingson, E., Faccioli, L., Fadeyev, V., Fakhouri, H. K., Fruchter, A. S., Gilbank, D. G., Gladders, M. D., Goldhaber, G., Gonzalez, A. H., Goobar, A., Gude, A., Hattori, T., Hoekstra, H., Hsiao, E., Huang, X., Ihara, Y., Jee, M. J., Johnston, D., Kashikawa, N., Koester, B., Konishi, K., Kowalski, M., Linder, E. V., Lubin, L., Melbourne, J., Meyers, J., Morokuma, T., Munshi, F., Mullis, C., Oda, T., Panagia, N., Perlmutter, S., Postman, M., Pritchard, T., Rhodes, J., Ripoche, P., Rosati, P., Schlegel, D. J., Spadafora, A., Stanford, S. A., Stanishev, V., Stern, D., Strovink, M., Takanashi, N., Tokita, K., Wagner, M., Wang, L., Yasuda, N., Yee, H. K. C., & Supernova Cosmology Project, T. (2012). The Hubble Space Telescope Cluster Supernova Survey. V. Improving the Dark-energy Constraints above z > 1 and Building an Early-type-hosted Supernova Sample. ApJ, 746, 85.

Tasse, C. (2014). Nonlinear Kalman filters for calibration in radio interferometry. A&A, 566, A127.

Tasse, C., Le Borgne, D., Röttgering, H., Best, P. N., Pierre, M., & Rocca-Volmerange, B. (2008). Radio-loud AGN in the XMM-LSS field. I. Optical identification and sample selection. A&A, 490, 879–891.

Tasse, C., van der Tol, S., van Zwieten, J., van Diepen, G., & Bhatnagar, S. (2013). Applying full polarization A-Projection to very wide field of view instruments: An imager for LOFAR. A&A, 553, A105.

Taylor, A. R., Bhatnagar, S., Condon, J., Green, D. A., Stil, J. M., Jagannathan, P., Kantharia, N., Kothes, R., Perley, R., Wall, J., & Willis, T. (2014). The Deep Full-Stokes Radio Sky. *ArXiv e-prints*.

Taylor, G. B., Carilli, C. L., & Perley, R. A., Eds. (1999). *Synthesis Imaging in Radio Astronomy II*, volume 180 of *Astronomical Society of the Pacific Conference Series*.

Taylor, J. H., Hulse, R. A., Margon, B., Davidsen, A., Mason, K., Sanford, P., Liller, W., Bernacca, P. L., Ciatti, F., John, R. S., Regener, V. H., Papaliolios, C., Pennypacker, C., Canizares, C., McClintock, J., Jones, B., Graham, D., & Wielebinski, R. (1974). Binary pulsar. IAU Circ., 2704.

Tripp, R. (1997). Using distant type IA supernovae to measure the cosmological expansion parameters. A&A, 325, 871–876.

Tyson, J. A., Kochanski, G. P., & Dell'Antonio, I. P. (1998). Detailed mass map of cl 0024+1654 from strong lensing. *The Astrophysical Journal Letters*, 498(2), L107.

van Haarlem, M. P., Wise, M. W., Gunst, A. W., Heald, G., McKean, J. P., Hessels, J. W. T., de Bruyn, A. G., Nijboer, R., Swinbank, J., Fallows, R., Brentjens, M., Nelles, A., Beck, R., Falcke, H., Fender, R., Hörandel, J., Koopmans, L. V. E., Mann, G., Miley, G., Röttgering, H., Stappers, B. W., Wijers, R. A. M. J., Zaroubi, S., van den Akker, M., Alexov, A., Anderson, J., Anderson, K., van Ardenne, A., Arts, M., Asgekar, A., Avruch, I. M., Batejat, F., Bähren, L., Bell, M. E., Bell, M. R., van Bemmel, I., Bennema, P., Bentum, M. J., Bernardi, G., Best, P., Bîrzan, L., Bonafede, A., Boonstra, A.-J., Braun, R., Bregman, J., Breitling, F., van de Brink, R. H., Broderick, J., Broekema, P. C., Brouw, W. N., Brüggen, M., Butcher, H. R., van Cappellen, W., Ciardi, B., Coenen, T., Conway, J., Coolen, A., Corstanje, A., Damstra, S., Davies, O., Deller, A. T., Dettmar, R.-J., van Diepen, G., Dijkstra, K., Donker, P., Doorduin, A., Dromer, J., Drost, M., van Duin, A., Eislöffel, J., van Enst, J., Ferrari, C., Frieswijk, W., Gankema, H., Garrett, M. A., de Gasperin, F., Gerbers, M., de Geus, E., Grießmeier, J.-M., Grit, T., Gruppen, P., Hamaker, J. P., Hassall, T., Hoeft, M., Holties, H. A., Horneffer, A., van der Horst, A., van Houwelingen, A., Huijgen, A., Iacobelli, M., Intema, H., Jackson, N., Jelic, V., de Jong, A., Juette, E., Kant, D., Karastergiou, A., Koers, A., Kollen, H., Kondratiev, V. I., Kooistra, E., Koopman, Y., Koster, A., Kuniyoshi, M., Kramer, M., Kuper, G., Lambropoulos, P., Law, C., van Leeuwen, J., Lemaitre, J., Loose, M., Maat, P., Macario, G., Markoff, S., Masters, J., McFadden, R. A., McKay-Bukowski, D., Meijering, H., Meulman, H., Mevius, M., Middelberg, E., Millenaar, R., Miller-Jones, J. C. A., Mohan, R. N., Mol, J. D., Morawietz, J., Morganti, R., Mulcahy, D. D., Mulder, E., Munk, H., Nieuwenhuis, L., van Nieuwpoort, R., Noordam, J. E., Norden, M., Noutsos, A., Offringa, A. R., Olofsson, H., Omar, A., Orrú, E., Overeem, R., Paas, H., Pandey-Pommier, M., Pandey, V. N., Pizzo, R., Polatidis, A., Rafferty, D., Rawlings, S., Reich, W., de Reijer, J.-P., Reitsma, J., Renting, G. A., Riemers, P., Rol, E., Romein, J. W., Roosjen, J., Ruiter, M., Scaife, A., van der Schaaf, K., Scheers, B., Schellart, P., Schoenmakers, A., Schoonderbeek, G., Serylak, M., Shulevski, A., Sluman, J., Smirnov, O., Sobey, C., Spreeuw, H., Steinmetz, M., Sterks, C. G. M., Stiepel, H.-J., Stuurwold, K., Tagger, M., Tang, Y., Tasse, C., Thomas, I., Thoudam, S., Toribio, M. C., van der Tol, B., Usov, O., van Veelen, M., van der Veen, A.-J., ter Veen, S., Verbiest, J. P. W., Vermeulen, R., Vermaas, N., Vocks, C., Vogt, C., de Vos, M., van der Wal, E., van Weeren, R., Weggemans, H., Weltevrede, P., White, S., Wijnholds, S. J., Wilhelmsson, T., Wucknitz, O., Yatawatta, S., Zarka, P., Zensus, A., & van Zwieten, J. (2013). LOFAR: The LOw-Frequency ARray. A&A, 556, A2.

van Weeren, R. J., Williams, W. L., Hardcastle, M. J., Shimwell, T. W., Rafferty, D. A., Sabater, J., Heald, G., Sridhar, S. S., Dijkema, T. J., Brunetti, G., Brüggen, M., Andrade-Santos, F., Ogrean, G. A., Röttgering, H. J. A., Dawson, W. A., Forman, W. R., de Gasperin, F., Jones, C., Miley, G. K., Rudnick, L., Sarazin, C. L., Bonafede, A., Best, P. N., Bîrzan, L., Cassano, R., Chyży, K. T., Croston, J. H., Ensslin, T., Ferrari, C., Hoeft, M., Horellou, C., Jarvis, M. J., Kraft, R. P., Mevius, M., Intema, H. T., Murray, S. S., Orrú, E., Pizzo, R., Simionescu, A., Stroe, A., van der Tol, S., & White, G. J. (2016). LOFAR Facet Calibration. ApJS, 223, 2.

Vrbanec, D., Ciardi, B., Jelić, V., Jensen, H., Zaroubi, S., Fernandez, E. R., Ghosh, A., Iliev, I. T., Kakiichi, K., Koopmans, L. V. E., & Mellema, G. (2015). Predictions for the 21cm-galaxy cross-power spectrum observable with LOFAR and Subaru. *ArXiv e-prints*.

Waddington, I., Dunlop, J. S., Peacock, J. A., & Windhorst, R. A. (2001). The LBDS Hercules sample of mJy radio sources at 1.4 GHz - II. Redshift distribution, radio luminosity function, and the high-redshift cut-off. MNRAS, 328, 882–896.

Wakker, B. P. & Schwarz, U. J. (1988). The Multi-Resolution CLEAN and its application to the short-spacing problem in interferometry. A&A, 200, 312–322.

Watson, G. N. (1995). *A treatise on the theory of Bessel functions*. Cambridge university press.

Webb, S. (1991). *Measuring the Universe: The Cosmological Distance Ladder*. Springer-Verlag London.

Weinberg, D. H., Mortonson, M. J., Eisenstein, D. J., Hirata, C., Riess, A. G., & Rozo, E. (2013). Observational probes of cosmic acceleration. Phys. Rep., 530, 87–255.

Weinstock, R. (2012). *Calculus of Variations*. Dover Books on Mathematics. Dover Publications.

White, R. L., Becker, R. H., Helfand, D. J., & Gregg, M. D. (1997). A Catalog of 1.4 GHz Radio Sources from the FIRST Survey. ApJ, 475, 479–493.

Whittaker, L., Brown, M. L., & Battye, R. A. (2015). Separating weak lensing and intrinsic alignments using radio observations. MNRAS, 451, 383–399.

Williams, W. L., Calistro Rivera, G., Best, P. N., Hardcastle, M. J., Röttgering, H. J. A., Duncan, K. J., de Gasperin, F., Jarvis, M. J., Miley, G. K., Mahony, E. K., Morabito, L. K., Nisbet, D. M., Prandoni, I., Smith, D. J. B., Tasse, C., & White, G. J. (2018). LOFAR-Boötes: properties of high- and low-excitation radio galaxies at 0.5 < z < 2.0. MNRAS, 475, 3429–3452.

Williams, W. L., van Weeren, R. J., Röttgering, H. J. A., Best, P., Dijkema, T. J., de Gasperin, F., Hardcastle, M. J., Heald, G., Prandoni, I., Sabater, J., Shimwell, T. W., Tasse, C., van Bemmel, I. M., Brüggen, M., Brunetti, G., Conway, J. E., Enßlin, T., Engels, D., Falcke, H., Ferrari, C., Haverkorn, M., Jackson, N., Jarvis, M. J., Kapińska, A. D., Mahony, E. K., Miley, G. K., Morabito, L. K., Morganti, R., Orrú, E., Retana-Montenegro, E., Sridhar, S. S., Toribio, M. C., White, G. J., Wise, M. W., & Zwart, J. T. L. (2016). LOFAR 150-MHz observations of the Boötes field: catalogue and source counts. MNRAS, 460, 2385–2412.

Wilman, R. J., Miller, L., Jarvis, M. J., Mauch, T., Levrier, F., Abdalla, F. B., Rawlings, S., Klöckner, H.-R., Obreschkow, D., Olteanu, D., & Young, S. (2008). A semi-empirical simulation of the extragalactic radio continuum sky for next generation radio telescopes. *Monthly Notices of the Royal Astronomical Society*, 388(3), 1335–1348.

Wong, K. C., Suyu, S. H., Auger, M. W., Bonvin, V., Courbin, F., Fassnacht, C. D., Halkola, A., Rusu, C. E., Sluse, D., Sonnenfeld, A., Treu, T., Collett, T. E., Hilbert, S., Koopmans, L. V. E., Marshall, P. J., & Rumbaugh, N. (2017). H0licow - iv. lens mass model of he 0435-1223 and blind measurement of its time-delay distance for cosmology. *Monthly Notices of the Royal Astronomical Society*, 465(4), 4895.

Zuntz, J., Kacprzak, T., Voigt, L., Hirsch, M., Rowe, B., & Bridle, S. (2013). IM3SHAPE: a maximum likelihood galaxy shear measurement code for cosmic gravitational lensing. MNRAS, 434, 1604–1618.

Zwicky, F. (1937). On the Masses of Nebulae and of Clusters of Nebulae. ApJ, 86, 217.

Appendix A

Form UPR16

FORM UPR16

Research Ethics Review Checklist

Please include this completed form as an appendix to your thesis (see the Postgraduate Research Student Handbook for more information

Postgraduate Research Student (PGRS) Information				**Student ID:**	480979
PGRS Name:	Michael Tarr				
Department:	ICG	**First Supervisor:**	David Bacon		
Start Date: (or progression date for Prof Doc students)	**01/10/2013**				
Study Mode and Route:	Part-time ☐	MPhil ☐	MD ☐		
	Full-time ☒	PhD ☒	Professional Doctorate ☐		

Title of Thesis:	Techniques for Cosmological Analysis of Next Generation Low to Mid-Frequency Radio Data.
Thesis Word Count: (excluding ancillary data)	51,523

If you are unsure about any of the following, please contact the local representative on your Faculty Ethics Committee for advice. Please note that it is your responsibility to follow the University's Ethics Policy and any relevant University, academic or professional guidelines in the conduct of your study

Although the Ethics Committee may have given your study a favourable opinion, the final responsibility for the ethical conduct of this work lies with the researcher(s).

UKRIO Finished Research Checklist:
(If you would like to know more about the checklist, please see your Faculty or Departmental Ethics Committee rep or see the online version of the full checklist at: http://www.ukrio.org/what-we-do/code-of-practice-for-research/)

a) Have all of your research and findings been reported accurately, honestly and within a reasonable time frame?	YES ☒ NO ☐
b) Have all contributions to knowledge been acknowledged?	YES ☒ NO ☐
c) Have you complied with all agreements relating to intellectual property, publication and authorship?	YES ☒ NO ☐
d) Has your research data been retained in a secure and accessible form and will it remain so for the required duration?	YES ☒ NO ☐
e) Does your research comply with all legal, ethical, and contractual requirements?	YES ☒ NO ☐

Candidate Statement:

I have considered the ethical dimensions of the above named research project, and have successfully obtained the necessary ethical approval(s)

Ethical review number(s) from Faculty Ethics Committee (or from NRES/SCREC):	9EA7-DEC4-CED6-157A-F1AB-6566-056C-168E

If you have *not* submitted your work for ethical review, and/or you have answered 'No' to one or more of questions a) to e), please explain below why this is so:

Appendix B

Ethical Review

Certificate of Ethics Review

Project Title:	Techniques for cosmological analysis of next generation data in the MhZ regime.
User ID:	480979
Name:	michael colin tarr
Application Date:	05/11/2015 17:09:53

You must download your referral certificate, print a copy and keep it as a record of this review.

The FEC representative for the School of Engineering is Marco Bruni

It is your responsibility to follow the University Code of Practice on Ethical Standards and any Department/School or professional guidelines in the conduct of your study including relevant guidelines regarding health and safety of researchers including the following:

- University Policy
- Safety on Geological Fieldwork

It is also your responsibility to follow University guidance on Data Protection Policy:

- General guidance for all data protection issues
- University Data Protection Policy

SchoolOrDepartment: ICG
PrimaryRole: PostgraduateStudent
SupervisorName: David Bacon
HumanParticipants: No
PhysicalEcologicalDamage: No
HistoricalOrCulturalDamage: No
HarmToAnimal: No
HarmfulToThirdParties: No
OutputsPotentiallyAdaptedAndMisused: No
Confirmation-ConsideredDataUse: Confirmed
Confirmation-ConsideredImpactAndMitigationOfPontentialMisuse: Confirmed
Confirmation-ActingEthicallyAndHonestly: Confirmed

Supervisor Review

As supervisor, I will ensure that this work will be conducted in an ethical manner in line with the University Ethics Policy.

Supervisor signature:

Date:

Appendix C

Catalogue sample

Below is a sample of the first 9 of 4,754 rows in the final catalogue created and used in chapter 3.

PLATE	FIBERID	OBJID	RA_SDSS	DEC_SDSS	OBJTYPE	CLASS	SUBCLASS
7562	1	(4679, 301, 1, 79, 325)	242.672	54.176	NA	GALAXY	STARBURST
7562	2	(4679, 301, 1, 79, 309)	242.716	54.231	NA	GALAXY	STARFORMING
7562	3	(4011, 301, 2, 75, 495)	242.892	54.215	NA	GALAXY	
7562	4	(4679, 301, 1, 79, 336)	242.617	54.150	NA	GALAXY	STARFORMING
7562	5	(4679, 301, 1, 80, 360)	242.853	54.230	NA	QSO	
7562	6	(4679, 301, 1, 80, 218)	242.824	54.143	NA	GALAXY	
7562	7	(4679, 301, 1, 80, 1439)	242.700	54.167	NA	QSO	
7562	8	(4011, 301, 2, 74, 933)	242.913	54.250	NA	GALAXY	
7562	9	(4679, 301, 1, 79, 1203)	242.622	54.168	NA	QSO	

...

SOURCETYPE	Z	Z_ERR	RCHI2	ZWARNING	SPEC1_G	SPEC1_R	SPEC1_I
ELAIS_N1_GMRT_GARN	0.2273	1.4499E-5	3.0208	0	21.4178	69.7417	54.1831
ELAIS_N1_GMRT_TAYLOR	0.0634	8.4446E-6	2.0018	0	21.4178	69.7417	54.1831
ELAIS_N1_GMRT_TAYLOR	0.2935	5.5908E-5	1.9839	0	21.4178	69.7417	54.1831
ELAIS_N1_GMRT_TAYLOR	0.2668	2.8540E-5	2.1935	0	21.4178	69.7417	54.1831
ELAIS_N1_GMRT_TAYLOR	2.2177	7.2118E-4	2.2127	0	21.4178	69.7417	54.1831
ELAIS_N1_GMRT_TAYLOR	0.0631	2.9196E-5	1.8625	0	21.4178	69.7417	54.1831
ELAIS_N1_GMRT_TAYLOR	6.5447	7.8872E-4	1.9859	4	21.4178	69.7417	54.1831
ELAIS_N1_GMRT_TAYLOR	0.6819	6.7622E-5	1.9478	0	21.4178	69.7417	54.1831
ELAIS_N1_GMRT_TAYLOR	0.8921	0.0239	1.8975	4	21.4178	69.7417	54.1831

...

SPEC2_G	SPEC2_R	SPEC2_I	SN_MEDIAN	SN_MEDIAN_ALL
22.8466	66.7091	44.5842	(3.4489, 12.6289, 16.5377, 18.7260, 14.3221)	15.6977
22.8466	66.7091	44.5842	(5.0193, 23.5378, 32.9386, 37.9620, 32.7787)	29.7295
22.8466	66.7091	44.5842	(0.2176, 1.6329, 3.2364, 4.1682, 3.8285)	2.7280
22.8466	66.7091	44.5842	(0.7692, 4.0730, 8.7305, 13.2010, 11.8439)	7.7951
22.8466	66.7091	44.5842	(1.7680, 5.8288, 4.9705, 4.7239, 3.8205)	4.9461
22.8466	66.7091	44.5842	(1.2374, 6.0577, 7.9359, 9.7762, 7.8775)	7.2340
22.8466	66.7091	44.5842	(0.2329, 0.5009, 0.6744, 1.3584, 1.3289)	0.7485
22.8466	66.7091	44.5842	(0.2866, 1.2882, 1.8085, 3.2942, 3.0127)	1.8194
22.8466	66.7091	44.5842	(0.3159, 0.6514, 1.2550, 2.5504, 2.5131)	1.2711

...

THROUGHPUT	SPECOBJID	MAG	XFOCAL	YFOCAL
22496	8514055484490839040	(21.0157, 20.1602, 19.5411, 19.1608, 19.0283)	289.82307	-66.063391
9263	8514055759368745984	(20.7307, 19.2033, 18.4434, 18.0221, 17.7052)	295.0358	-53.789525
282246	8514056034246652928	(23.87, 23.0909, 21.9004, 21.6587, 21.2116)	317.85234	-56.573507
13481	8514056309124559872	(23.5467, 21.8594, 20.684, 19.9961, 19.5377)	282.96523	-71.874449
587	8514056584002466816	(21.774, 21.488, 21.4103, 21.4432, 20.9497)	312.60659	-53.526114
10733	8514056858880373760	(22.665, 21.5083, 20.7331, 20.5374, 20.3089)	309.63221	-72.468989
11839	8514057133758280704	(24.3571, 24.3739, 23.7874, 22.9613, 22.1575)	293.4808	-67.776261
11546	8514057408636187648	(23.3142, 23.7815, 22.6532, 21.967, 21.3537)	320.15422	-48.750214
19556	8514057683514094592	(24.0226, 24.453, 23.3662, 22.1882, 21.9048)	283.50415	-67.81215

...

ANCILLARY_TARGET2	RUN	RERUN	CAMCOL	FIELD	ID	survey	survey_code	Mclass	Multid
1152921504606846976	4679	301	1	79	325	GMRT	3	S	-99
2305843009213693952	4679	301	1	79	309	Taylor_gmrt	4	S	-99
2305843009213693952	4011	301	2	75	495	Taylor_gmrt	4	S	-99
2305843009213693952	4679	301	1	79	336	Taylor_gmrt	4	S	-99
2305843009213693952	4679	301	1	80	360	Taylor_gmrt	4	S	-99
2305843009213693952	4679	301	1	80	218	Taylor_gmrt	4	S	-99
2305843009213693952	4679	301	1	80	1439	Taylor_gmrt	4	S	-99
2305843009213693952	4011	301	2	74	933	Taylor_gmrt	4	S	-99
0	4011	301	2	74	265				

...

surveyid	samplecode	RA_radio	DEC_radio	dr9_OBJID	size
GMRTEN1_J161041.1+541031	gold	242.67162	54.17553	1237665569290780997	1.8275
2138	gold	242.7159	54.231	1237665569290780981	2.3337
2552	gold	242.89306	54.2148	1237662700789236207	2.6855
1893	gold	242.61712	54.1497	1237665569290781008	1.6207
2456	gold	242.85279	54.2295	1237665569290846568	1.1725
2397	silver	242.82409	54.1428	1237665569290846426	0.7518
2112	bronze	242.69992	54.1672	1237665569290847647	1.9959
2599	gold	242.91243	54.2504	1237662700789171109	0.0

...

flags	flags_u	flags_g	flags_r	flags_i	flags_z
	2252075094278160	2252074960060432	6755674587430928	2252074968584208	2252075094278160
	2252074968584208	2252074960060432	6755674587566096	2252074960060432	2252074960191504
	20406935945879569	2269667146203152	6755674587529232	2252074960027664	2252074960060432
	2322168969068817	158604955123728	6755674587566608	2252074960060432	275280592912
	158604955123712	275146375168	4503874773745664	275146375168	275280592896
	211106643738897	142639377	76561193967321104	134250513	134381585
	18014398652121345	9007199397380353	9007199397380353	49539871324274688	140737622738177
	47287796229996817	2251800224727312	51791670861103120	2252074960060432	2252075094410512

...

modelMag_u	modelMag_g	modelMag_r	modelMag_i	modelMag_z	Rband_flux	Radio_flux
20.7838	19.7741	18.8391	18.3426	18.1467	1.0577E-4	0.0013
18.5484	17.4635	16.9340	16.5128	16.3467	6.1148E-4	0.0020
24.6106	21.1864	20.0083	19.4684	18.8051	3.6031E-5	0.0033
23.1526	20.8080	19.0173	18.3923	17.9312	8.9759E-5	0.0019
22.1290	21.0500	19.4816	18.9500	18.4921	5.8525E-5	0.0023
24.6194	23.1390	22.1111	20.9980	20.3281	5.1946E-6	0.0307
22.2642	21.2158	21.0504	20.6872	20.2063	1.3798E-5	nan
23.4206	22.0993	20.5135	19.5972	19.2640	2.2626E-5	0.0030
22.6165	20.1437	18.4251	17.5761	17.1635	1.5486E-4	7.6854E-4

...

modelMagErr_u	modelMagErr_g	modelMagErr_r	modelMagErr_i	modelMagErr_z	fracDeV_u	fracDeV_g
0.0445	0.0110	0.0084	0.0083	0.0219	0.9	0.9
0.0304	0.0055	0.0042	0.0042	0.0083	0.7	0.9
0.6700	0.1023	0.0507	0.0669	0.1533	1.0	0.0
0.8195	0.0377	0.0184	0.0148	0.0331	0.0	0.0
0.0828	0.0273	0.0299	0.0376	0.0977	1.0	1.0
0.6168	0.5547	0.0848	0.3466	0.4834	1.0	1.0
2.0318	0.4974	0.3883	0.1761	0.3707	0.0	0.0
0.5011	0.1778	0.1032	0.0817	0.1810	1.0	1.0
0.9377	0.4195	0.1690	0.0951	0.3247	1.0	1.0

...

fracDeV_r	fracDeV_i	fracDeV_z	fieldID	Type	Line	Linewl	Restwl	Redshift	Confidence
0.924	0.866	0.846	1237665569290780672	e	Hbeta	5964.135	4862.68	0.2265	1
0.993	1.0	1.0	1237665569290780672	e	Halpha	6989.057	6564.61	0.0647	1
0.0	0.0	0.0	1237662700789235712	e	OIII	9322.39	5008.2	0.8614	3
0.916	0.04	0.459	1237665569290780672	e	Halpha	8319.275	6564.61	0.2673	1
1.0	0.0	1.0	1237665569290846208	q	CIV	4971.923	1549.06	2.2096	1
0.0	1.0	1.0	1237665569290846208	g	CaHK	4186.877	3934.78	0.0641	2
0.0	0.578	1.0	1237665569290846208	?	Unknown	5266.316	0.0	0.0	3
0.0	0.0	0.154	1237662700789170176	?	Unknown	7119.898	0.0	0.0	3
1.0	0.935	1.0	1237665569290780672	?	Unknown	5866.004	0.0	0.0	3

Appendix D

Taylor Expansion of Lensing model

This appendix contains the derivation of the Taylor expansion of Lensing Jacobian that was required in section 5.1.1.

D.1 First order terms

Starting with equation 5.16 that describes the initial sky model, we need to find the partial derivatives of I w.r.t γ_1 and γ_2, where,

$$I(\mathbf{x}, \gamma) = f_g(|\mathscr{A}_g \mathbf{x}|)$$

and $\mathscr{A}$ is given by,

$$\mathscr{A}(\gamma) = \begin{pmatrix} 1-\gamma_1 & -\gamma_2 \\ -\gamma_2 & 1+\gamma_1 \end{pmatrix}$$

.

Considering the interior function we have,

$$A = |\mathscr{A}\mathbf{x}| = ((x-\gamma_1 x-\gamma_2 x)^2 + (y+\gamma_1 y-\gamma_2 x)^2)^{\frac{1}{2}} \tag{D.1}$$

$$A_{\gamma_1} = \frac{1}{2}(-2x(x-\gamma_1 x-\gamma_2 y)+2y(y+\gamma_1 y-\gamma_2 x)) \\ ((x-\gamma_1 x-\gamma_2 x)^2 + (y+\gamma_1 y-\gamma_2 x)^2)^{-1/2} \tag{D.2}$$

$$A_{\gamma_2} = \frac{1}{2}(-2y(x-\gamma_1 x-\gamma_2 y)-2x(y+\gamma_1 y-\gamma_2 x)) \\ ((x-\gamma_1 x-\gamma_2 x)^2 + (y+\gamma_1 y-\gamma_2 x)^2)^{-1/2}. \tag{D.3}$$

Providing,

$$A_{\gamma_1}(\gamma=0) = \frac{y^2-x^2}{|\mathbf{x}|} \tag{D.4}$$

$$A_{\gamma_2}(\gamma=0) = \frac{-2xy}{|\mathbf{x}|} \tag{D.5}$$

Hence, we can now use this result, substituting into equation 5.7 in section 5.1.1.

D.2 Second Order Terms

Extending the first order terms above to second order is also interesting. I define the second order terms of the Taylor expansion as I_2 where,

$$I_2 = \frac{1}{2}(\gamma_1^2 I_{\gamma_1\gamma_1}(\gamma=0) + 2\gamma_1\gamma_2 I_{\gamma_1\gamma_2}(\gamma=0) + \gamma_2^2 I_{\gamma_2\gamma_2}(\gamma=0))$$

.

This gives,

$$I_{\gamma_i\gamma_j} = \frac{\partial^2 f}{\partial\gamma_i\partial\gamma_j} = \frac{\partial}{\partial\gamma_j}\left(\frac{\partial f}{\partial\gamma_i}\right) \tag{D.6}$$

$$= \frac{\partial}{\partial\gamma_j}\left(\frac{\partial f}{\partial A}\frac{\partial A}{\partial\gamma_i}\right) \tag{D.7}$$

$$= \frac{\partial f'}{\partial\gamma_j}\frac{\partial A}{\partial\gamma_i} + f'\frac{\partial^2 A}{\partial\gamma_i\partial\gamma_j} \tag{D.8}$$

$$= \frac{\partial f'}{\partial A}\frac{\partial A}{\partial\gamma_j}\frac{\partial A}{\partial\gamma_i} + f'\frac{\partial^2 A}{\partial\gamma_i\partial\gamma_j} \tag{D.9}$$

$$= f''\frac{\partial A}{\partial\gamma_j}\frac{\partial A}{\partial\gamma_i} + f'\frac{\partial^2 A}{\partial\gamma_i\partial\gamma_j} \tag{D.10}$$

Hence

$$I_{\gamma_1\gamma_1} = f''A_{\gamma_1}^2 + f'A_{\gamma_1\gamma_1} \tag{D.11}$$

$$I_{\gamma_2\gamma_2} = f''A_{\gamma_2}^2 + f'A_{\gamma_2\gamma_2} \tag{D.12}$$

$$I_{\gamma_1\gamma_2} = f''A_{\gamma_1}A_{\gamma_2} + f'A_{\gamma_1\gamma_2} \tag{D.13}$$

From Appendix D.1,

$$A_{\gamma_1} = (-x^2 + y^2 + \gamma_1 x^2 + \gamma_1 y^2)A^{-1} \tag{D.14}$$
$$A_{\gamma_2} = (-2xy + \gamma_2 y^2 + \gamma_2 x^2)A^{-1} \tag{D.15}$$

Which gives

$$A_{\gamma_1\gamma_1} = (x^2 + y^2)A^{-1} - (-x^2 + y^2 + \gamma_1 x^2 + \gamma_1 y^2)^2 A^{-3} \tag{D.16}$$
$$A_{\gamma_2\gamma_2} = (x^2 + y^2)A^{-1} - (-2xy + \gamma_2 x^2 + \gamma_2 y^2)^2 A^{-3} \tag{D.17}$$
$$A_{\gamma_1\gamma_2} = (-x^2 + y^2 + \gamma_1 x^2 + \gamma_1 y^2)(2xy - \gamma_2 x^2 - \gamma_2 y^2)A^{-3} \tag{D.18}$$
$$A_{\gamma_1\gamma_1}(\gamma = 0) = \frac{x^2 + y^2}{|\mathbf{x}|} - \frac{(-x^2 + y^2)^2}{|\mathbf{x}|^3} = |\mathbf{x}| - \frac{(-x^2 + y^2)^2}{|\mathbf{x}|^3} \tag{D.19}$$
$$A_{\gamma_2\gamma_2}(\gamma = 0) = \frac{x^2 + y^2}{|\mathbf{x}|} - \frac{(-2xy)^2}{|\mathbf{x}|^3} = |\mathbf{x}| - \frac{(2xy)^2}{|\mathbf{x}|^3} \tag{D.20}$$
$$A_{\gamma_1\gamma_2}(\gamma = 0) = \frac{-(-x^2 + y^2)(-2xy)}{|\mathbf{x}|^3} = \frac{2(-x^2 + y^2)xy}{|\mathbf{x}|^3} \tag{D.21}$$

Converting to polar coordinates,

$$A_{\gamma_1}(\gamma = 0) = -r\cos 2\theta \tag{D.22}$$
$$A_{\gamma_2}(\gamma = 0) = -r\sin 2\theta \tag{D.23}$$
$$A_{\gamma_1\gamma_1}(\gamma = 0) = r\sin^2 2\theta \tag{D.24}$$
$$A_{\gamma_2\gamma_2}(\gamma = 0) = r\cos^2 2\theta \tag{D.25}$$
$$A_{\gamma_1\gamma_2}(\gamma = 0) = -\frac{1}{2} r\sin 4\theta \tag{D.26}$$

which gives

$$I_{\gamma_1\gamma_1} = f'' r^2 \cos^2 2\theta + f' r \sin^2 2\theta \tag{D.27}$$
$$I_{\gamma_2\gamma_2} = f'' r^2 \sin^2 2\theta + f' r \cos^2 2\theta \tag{D.28}$$
$$I_{\gamma_1\gamma_2} = f'' r^2 \cos 2\theta \sin 2\theta - \frac{1}{2} f' r \sin 4\theta \tag{D.29}$$

Hence,

$$I_2 = \frac{1}{2}\left(f''r^2\left(\gamma_1^2\cos^2 2\theta + 2\gamma_1\gamma_2\cos 2\theta\sin 2\theta + \gamma_2^2\sin^2 2\theta\right)\right. \tag{D.30}$$

$$\left. + f'r\left(\gamma_1^2\sin^2 2\theta - 2\gamma_1\gamma_2\sin 2\theta\cos 2\theta + \gamma_2^2\cos^2 2\theta\right)\right) \tag{D.31}$$

$$= \frac{1}{2}\left(f''r^2(\gamma_1\cos 2\theta + \gamma_2\sin 2\theta)^2 + f'r(\gamma_1\sin 2\theta - \gamma_2\cos 2\theta)^2\right) \tag{D.32}$$

So we see a repetition of the same lensing kernel $\mathbb{K}$ with higher order derivatives of the source profile. This is fascinating and reveals that a similar approach to FILM could be used to detect these distortions, given the requite data quality.